THE LAT ANALYSIS TOOLS

Monday, December 2, 2002

Table of Contents

1 Introduction

This document describes the LAT analysis tools that will be used to derive astrophysical results from the LAT data. These tools are a major component of the "standard analysis environment" (SAE). The LAT data is assumed to have already undergone Level 1 processing, where the events are reconstructed, and the events most likely to be photons are identified. The gamma-ray burst tools will also be capable of analyzing GBM data, as well as data from other missions (e.g., Swift).

The level of detail in this document will increase with time. Currently the goal is to describe the function of each tool, the planned input and output, and rudiments of the algorithms for some of the tools. Eventually, this document will grow into a formal requirements document.

2 Data Formats

We define the communication between tools in terms of FITS files. In some environments the communication may take other forms (e.g., data passed in memory in an environment such as GAUDI), but all communication modes should contain the information in these file definitions, and all the tools should be capable of actually reading and writing these files. Many of the files have similar contents, and thus will have the same format. Therefore, we define a number of standard data file formats. Formatting the files using FITS will permit the use of FTOOLS to examine and manipulate the contents of these files. In the descriptions below we point out some of the features of the tables and headers.

In the input and output subsections of each tool description the FITS file type is identified. Conversely, in the file descriptions below the tools that use these file types are identified.

2.1 FT1—Event Data

Contents of FITS file:
Header—contains information about operations (e.g., cuts) performed on the event list. This information is relevant to the calculation of the exposure. If the barycentric time and pulsar phase are nonzero, then the pulsar name must be included. The header also indicates whether the data are real or simulated.
Extension table—the format is the same as a D1 table. Two additional columns are added: barycentric time and pulsar phase.

	Column Name	Units
1	Energy of event	GeV
2	1-sigma uncertainty of energy	GeV
3	RA, J2000	deg
4	Dec, J2000	deg
5	Localization uncertainty, est. 1-sigma radius	deg
6	Time (Mission Elapsed Time)	S
7	Conversion layer	dimensionless
8	Number of SSD hits	dimensionless
9	Number of tracker hits NOT reconstructed	dimensionless
10	Conversion point, (x,y,z)	M
11	Reconstruction trajectory-event (dir. cosine)	dimensionless
12	Reconstruction trajectory-secondary 1 (dir. cosine)	dimensionless
13	Reconstruction trajectory secondary 2 (dir. cosine)	dimensionless
14	Secondary energies	GeV
15	ACD tiles hit (bit flags)	dimensionless

16	Quality_Parm--quality parameter for fitted trajectory	dimensionless
17	Data_Quality--Overall status of event	dimensionless
18	Deadtime accumulated since start of mission	S
19	Instrument mode (slew, diagnostic, ...)	dimensionless
20	TKR, CAL-only flag--2bit for CAL, TKR or Both	dimensionless
21	Zenith angle	deg
22	Earth azimuth angle (from north to east)	deg
23	S/C position from earth center, (x,y,z)*	km
24	Angle from sun	deg
25	Angle from moon	deg
26	Ground point--lat.	deg
27	Ground point--lon.	deg
28	Barycenter arrival time of event	s
29	Offset of Solar System Barycenter from S/C (x,y,z)	s (light travel time)
30	McIlwain B	gauss
31	McIlwain L	Earth radii
32	Geomagnetic Lat.	deg
33	Geomagnetic Long.	deg
34	Recon. Version	dimensionless
35	Calib. table versions	dimensionless
36	Multiple event reconstruction	dimensionless
37	Reconstruction number within event	dimensionless
38	Original Event ID	dimensionless
39	Barycentric time	S
40	Pulsar phase	dimensionless

Uses:
Input to:
- U10—Arrival time correction
- A4—Pulsar period search
- U12—Pulsar phase assign
- A3—Pulsar profiles
- U4—Exposure calculation
- A1--Likelihood
- U6—Mapgeneration
- A5—GRB event binning
- A9—GRB unbinned spectral analysis
- A10—GRB spectral-temporal modeling
- A7—GRB temporal analysis
- U13—GRB visualization
- U2—Data subselection

Output from:
- U1—Data extract
- U2—Data subselection

- U10—Arrival time correction
- U12—Pulsar phase assign
- O2—Observation simulator

Issues:
1. Do we want/need to carry around all these quantities?
2. Should the pulsar quantities be additional columns? Should they always be present or can the table be extended easily? Should they be in an additional extension table?

2.2 FT2—Pointing History

Contents of FITS file:
Header—indicates whether the pointing history is real or simulated.
Extension table—format same as D2 table.

	Contents	Units
1	starting time of interval (Mission Elapsed Time)	s
2	ending time of interval (Mission Elapsed Time)	s
3	position of S/C at start of interval (x,y,z inertial coordinates)	km
4	viewing direction at start (LAT +z axis), 2 angles	dimensionless
5	orientation at start (LAT +x axis), 2 angles	dimensionless
6	zenith direction at start, 2 angles	dimensionless
7	LAT operation mode	dimensionless
8	livetime	s
9	SAA flag	dimensionless
10	S/C longitude	deg
11	S/C longitude	deg
12	S/C altitude	km
13	direction of the sun, 2 angles	Deg
14	direction of the moon, 2 angles	

Uses:
Input to:
- U4—Exposure calculation
- U14—GRB LAT DRM generator
- O2—Observation simulator

Output from:
- U3—Pointing/livetime extractor
- O1—Pointing/livetime simulator

2.3 FT3—XSPEC Format PHA2

These files result from binning event data from gamma-ray bursts. Their format is that of the standard XSPEC input file.

Uses:
Input to:
- A6—GRB rebinning
- U14—GRB LAT DRM generator
- A8—GRB binned spectral analysis (XSPEC)
- U13—GRB visualization
- A7—GRB temporal analysis

Output from:
- A5—GRB event binning
- A6—GRB event rebinning
- U14—GRB LAT DRM generator

2.4 FT4, FT5—XSPEC Format ARF and RMF

These files are required for fitting binned 1D burst spectra. Their format is that of the standard XSPEC input files. The ARF contains the effective area part of the DRM, and the RMF the energy redistribution matrix.

Uses:
Input to:
- A8—GRB binned spectral analysis (XSPEC)
- U13—GRB visualization

Output from:
- U14—GRB LAT DRM generator

2.5 FT6—Spectral Models

These files result from fitting sequences of burst spectra by XSPEC. The format may be based on XSPEC XCM files, which contain the model fit for only one spectrum per file. These files should list the PHA2, ARF, RMF, and background files, identify the model function (e.g., power law), and then provide the fitted parameters.

Uses:
Input to:
- U13—GRB visualization

Output from:

- A8—GRB binned spectral analysis (XSPEC)
- A9—GRB unbinned spectral analysis

2.6 FT7—Exposure

These files contain the factors of the likelihood that need be calculated only once for a given analysis. These factors usually have units of area*time, i.e, exposure. The precise form of these factors will depend on the formulation of the likelihood used in A1.

Uses:
Input to:
- A1—Likelihood analysis
- U6—Map generator
- A9—GRB unbinned spectral analysis
- A10—GRB spectral-temporal modeling

Output from:
- U4—Exposure calculation

2.7 FT8—Source Definition

These files contain the description of a region of the sky.

Uses:
Input to:
- A1—Likelihood
- O2—Observation simulator
- U7—Source model definition

Output from:
- A1—Likelihood
- U7—Source model definition

2.8 Other files:

- Pulsar ephemeris—From U11 (Ephemeris extract) to U12 (Pulsar phase assign)
- Energy binning grid—Input to A5 (GRB event binning)
- Time binning grid—Input to A5 (GRB event binning)

3 Databases

This section lists the databases that are relevant to the standard analysis environment (SAE). Other databases will be maintained by the LAT team or the SSC but are not considered part of the SAE.

In most cases these databases will be maintained in a HEASARC-standard form: data in FITS files with metadata pointing to these FITS files maintained in a relational database. In some cases a special operational form might be created (e.g., for rapid access), but ultimately all the databases will be converted to HEASARC-standard format for archiving with the HEASARC.

The data necessary for the calculation of the instrument response functions will be maintained in a CALDB directory structure.

D1. Event summary (Level 1) database

Date: 13 September 2002 improved some comments in the Existing counterparts section; 9 Sept 2002 (draft v3,MH removed barycenter offset; 28 August 2002 SD removed barycenter arrival time of gamma ray, corrected event rate from 25 to 30 Hz, record size from 250 to 200 bytes; 7 May 2002, now includes sun and moon angles); 12/1/02 DLB

Function

This database will house all the reconstructed events from the LAT and carry all the information necessary for higher analysis. It is not directly accessed by the analysis tools. It receives queries from, and passes data to, the Data Extractor (U1).

Contents

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-002 (and updated to include the angles of the sun and moon – see U4), are as follows:

	Column Name	Units
1	Energy of event	GeV
2	1-sigma uncertainty of energy	GeV
3	RA, J2000	deg
4	Dec, J2000	deg
5	Localization uncertainty, est. 1-sigma radius	deg
6	Time (Mission Elapsed Time)	s
7	Conversion layer	dimensionless
8	Number of SSD hits	dimensionless
9	Number of tracker hits NOT reconstructed	dimensionless
10	Conversion point, (x,y,z)	m
11	Reconstruction trajectory-event (dir. cosine)	dimensionless
12	Reconstruction trajectory-secondary 1 (dir. cosine)	dimensionless
13	Reconstruction trajectory secondary 2 (dir. cosine)	dimensionless
14	Secondary energies	GeV
15	ACD tiles hit (bit flags)	dimensionless
16	Quality_Parm--quality parameter for fitted trajectory	dimensionless
17	Data_Quality--Overall status of event	dimensionless
18	Deadtime accumulated since start of mission	s
19	Instrument mode (slew, diagnostic, ...)	dimensionless
20	TKR, CAL-only flag--2bit for CAL, TKR or Both	dimensionless
21	Zenith angle	deg
22	Earth azimuth angle (from north to east)	deg
23	S/C position from earth center, (x,y,z)*	km
24	Angle from sun	deg
25	Angle from moon	deg

26	Ground point--lat.	deg
27	Ground point--lon.	deg
28	McIlwain B	gauss
29	McIlwain L	Earth radii
30	Geomagnetic Lat.	deg
31	Geomagnetic Long.	deg
32	Recon. Version	dimensionless
33	Calib. table versions	dimensionless
34	Multiple event reconstruction	dimensionless
35	Reconstruction number within event	dimensionless
36	Original Event ID	dimensionless

* x-axis is direction RA,DEC = 0,0, z-axis is north, y-axis defined for earth-centered right-hand coordinate system, J2000

Also from the DPWG report:
The reconstructed trajectory of the photon (item 13) is in instrument coordinates, and so specifies the inclination angle and azimuth of the photon.
Instrument modes (item 19) should probably be defined in a second extension.
Perhaps multi-gamma events should just have a 'primary' gamma ray defined here, plus a flag set to indicate that special processing should be done.
Note that >1 photon can be claimed by Reconstruction (may translate into multiple Event Summary entries, with number_of_photons entry)
Quality flags above are intended to be representative of the background rejection/PSF enhancement cut parameters

Performance requirements

See LAT Event Summary Database Requirements Document

Host environment

Database server system

Existing counterparts

Nothing directly applicable. Spatial indexing methods (hierarchical tessellations of the sky) for facilitating access to data by region of the sky have been designed for other astronomical applications. Two widely-used indexing schemes (HTM and HEALpix) are well documented and some software is available for efficient coordinate-spatial index conversion is available. Owing to the scanning mode of operation of the LAT, which implies that data for a given region of the sky are spread out in arrival time, and the nature of the Level 1 database (relatively static, with most access being read-only), sequential searches of flat files offer performance competitive to relational database searches.

Each entry requires roughly 200 bytes of storage, and with a telemetry-limited event rate of 30 Hz (including cosmic rays), this yields about 200 Gb of data a year.

Data retrieval times from 2-dimensional spatial queries are the most relevant performance statistic. We are currently benchmarking performance to determine these and other performance measures for potential database architectures. One enhancement we are likely to make is to separate the gamma rays into a their own database since they will be the primary focus of the analysis. (Cosmic-ray events in the telemetry will primarily be used by the LAT team for low-level calibration.) Rapid retrieval of data from a 20 Gb/year photon database should not be problematic. Other issues being considered are ease of maintenance, ease of updates (due to reprocessing), and ease of backup. We expect to choose the database architecture by the end of 2002.

Open issues for definition or implementation

1. Some input parameters for the Data Extractor (U3) might not be passed on to the Event summary database, depending on whether the Data Extractor does any ‘post processing’ on retrieved data and on the implementation of U2. This is probably neither here nor there as far as the requirements summaries go, however.

D2. Pointing, livetime, and mode history

Date: 13 September 2002 (draft v7 fixed reference to Pointing extractor); 7 May 2002 (draft v2, added sun and moon positions, v7, added); 12/01/02 DLB

Function

This is the database of pointing and observation history that is needed to calculate exposures. It contains information about the orientation, location, and operation of the LAT for regular time intervals, ~30 s. It is not directly accessed by the analysis tools. Instead, it receives queries from, and passes data to, the Pointing/livetime history extractor (U3). The positions of the sun and moon are included here solely to facilitate cuts on their positions in the generation of exposure. They are both gamma-ray sources (the sun impulsively) and both of course shadow sources they pass in front of.

Contents

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-005 and augmented here to include the SAA flag and positions of the sun and moon, are as follows:

	Contents	Units
1	starting time of interval (Mission Elapsed Time)	s
2	ending time of interval (Mission Elapsed Time)	s
3	position of S/C at start of interval (x,y,z inertial coordinates)	km
4	viewing direction at start (LAT +z axis), 2 angles	dimensionless
5	orientation at start (LAT +x axis), 2 angles	dimensionless
6	zenith direction at start, 2 angles	dimensionless
7	LAT operation mode	dimensionless
8	livetime	s
9	SAA flag	dimensionless
10	S/C longitude	deg
11	S/C longitude	deg
12	S/C altitude	km
13	direction of the sun, 2 angles	deg
14	direction of the moon, 2 angles	deg

Performance requirements

Performance requirements will be considered in detail in the full requirements document. None of the requirements is likely to be challenging in terms of implementation of the database. See Science Tools Database Requirements Document

Host environment

Database server system

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

1. Should latitude, longitude, and altitude be translated into geomagnetic coordinates in anticipation of any need to generate exposure maps for different ranges of geomagnetic conditions? (This might be useful, e.g., if the background rejection is not extremely good.)

D3. Instrument response functions

Date: 6 Sept 2002 (draft v2, updated 15 Sept 2002); 12/1/02 DLB

Function

The GLAST CALDB is the directory and indexing structure that stores the calibration files required for producing the instrument response functions (IRFs) for a given analysis; the analysis tools will extract the relevant data from the CALDB structure. IRFs are the high-level specifications of the performance of the LAT. As is customary, the response will be decomposed into the point-spread, effective area, and energy redistribution functions. These response functions depend on photon energy, direction of incidence, plane of conversion in the LAT tracker, and event quality flags. They may also be expected to be functions of time, in the event that the instrumental response changes significantly, e.g., as a result of partial failure of one subsystem.

At a minimum the CALDB structure stores those FITS-format calibration files that are directly necessary for calculating the IRFs. Thus CALDB may include files with the PSF as a function of energy and inclination angle, energy redistribution matrices, etc. Different versions of the calibration files are retained, permitting users to replicate previous analyses that used earlier versions of the calibration files.

Thus the functions of the CALDB system are:
1. To store and archive calibration files.
2. To provide a naming convention and header structure for all calibration files.
3. To index the calibration files for software access based on FITS header keywords.
4. To update the calibration data independent of software updates, while maintaining configuration control.
5. To provide a traceable history of the database's calibration files by maintaining and identifying different versions.
6. To provide an interface between calibration files in a FITS-format understood by the IRF-generating software and IRF files with recognized FITS-formats (e.g., PHA files understood by XSPEC).

Contents

The exact set of instrumental parameters to be used for tabulating the response is still being evaluated. For example, the scientific return for tabulating the variation of energy resolution plane-by-plane within the tracker vs. tabulating it just for the front and back sections of the tracker, needs to be determined via analysis of simulations. (It should be noted that response functions for 'CAL-only' events where the gamma ray does not convert until it enters the calorimeter likely also will be needed.) The best choice of event quality flags will be established by study of the behavior of event reconstructions. It must be decided whether the response functions should be represented as parameterized functions or matrices of numbers.

If the functions are to be stored as matrices, the grid points must be spaced closely enough to allow linear interpolation. Here we present some rough estimates of the storage required, subject to revision. The photon energy values will necessarily be unequally spaced, perhaps equally spaced in log(energy). Approximately 200 energy values should be adequate. About 50 values of polar angle of incidence should suffice to cover the entire field of view. Equal spacing in the cosine of the angle may be the most useful choice. It is expected that there will be little variation in the response functions with azimuthal angle of incidence, but this must be checked with simulations. If azimuth is relevant, then at most 16 values should be needed. The number of event types will probably not exceed 10. Combining these sizes leads to an upper limit of 200 * 50 * 16 * 10 = 1.6 million floating point numbers for the effective area matrix. This is a very modest amount of storage, even if different tabulations are maintained for each plane of conversion.

The other response functions will require more storage. The energy redistribution function will have another index for the measured energy, with perhaps 50 values needed. This leads to a matrix size of 80 million numbers. If the point spread function is radially symmetric in the estimated angle, then it will require the same storage. It is likely that there will be a significant “fisheye” distortion at large inclinations, so a further factor of up to 40 may be needed. That would lead to an upper-limit storage requirement of 3.2 billion numbers per plane of conversion tabulated. Such a large object may become unwieldy, so it will be necessary to seek ways to factor the problem further.

The CALDB system defined by HEASARC can provide a workable implementation of the response functions, and that is the intended approach. Some of the standard CALDB file formats are suitable for the matrices described above. The HEASARC provides utilities for both remote access and local installation of CALDB files.

Performance requirements

Performance requirements will be considered in detail in the full requirements document. None of the requirements is likely to be challenging in terms of implementation of the database.

Host environment

At the SSC the GLAST CALDB will be incorporated within the HEASARC’s CALDB system, but other sites will have stand-alone CALDB structures modeled after the HEASARC’s system. User may download a subset of the CALDB containing only the desired version of the calibration files, or may access the CALDB remotely via the internet.

Existing counterparts

The basic structure and use of the CALDB has been specified by the HEASARC. Documentation for setting up and managing a HEASARC-compatible CALDB exists online at http://heasarc.gsfc.nasa.gov/docs/heasarc/caldb/caldb_doc.html, wherein numerous documents are available.

Open issues for definition or implementation

1. There is an area in which the needs of the LAT don't match up very well with CALDB. Most of the existing CALDB formats prefer that the measured energy be expressed in PHA channels. The energy of each LAT photon is estimated by combining information from several calorimeter PHAs and from many tracker hits. Variations in detector gains and in the list of dead tracker strips will be absorbed in the algorithm. Of course, this will require a new category of calibration data objects for this information.

D4. Pulsar ephemerides

Date: 13 September 2002 (draft v8, added link to req. document); 9 September 2002 (draft v4, edited by M. Hirayama); 12/1/02 DLB

Function

This is the pulsar timing information to be maintained during the LAT mission for assigning pulsar phases to gamma rays. A user can access it by using the Pulsar Ephemeris extractor (U11) or the Catalog Access tool (U9). The pulsar timing information mainly originates from radio timing observations, but this database will contain radio-quiet pulsars (e.g. Geminga pulsar).

Contents

The contents of the database are itemized below. They are defined based on the pulsar timing files used for EGRET for familiarity to the user community. For generality and consistency with format provided by pulsar observers, times in this file should be specified in MJD rather than Mission Elapsed Time. The second table below contains the additional information required for binary pulsars.

Spin parameters

Item	Contents	Units
1	Pulsar name	dimensionless
2	Right Ascension (J2000)	deg
3	Declination (J2000)	deg
4	Start of interval of validity for timing info (MJD)	days
5	End of interval of validity (MJD)	days
6	Infinite-frequency geocentric UTC arrival time of a pulse (MJD)	days
7	Pulsar rotation frequency (*)	Hz
8	First derivative of pulsar frequency (*)	Hz^2
9	Second derivative of pulsar frequency (*)	Hz^3
10	Root-mean-square radio timing residual (periods)	dimensionless
11	Source of timing information	dimensionless
12	Flag for binary pulsars	dimensionless

(*) The integer part of item 6 is the barycentric (TDB) epoch of items 7-10. The integer part and the fraction part should be stored separately, because standard double precision is not sufficient to keep the required precision for the sum.

Orbital parameters: for binary pulsars only

Item	Contents	Units
13	Orbital period	s

14	Projected semi-major axis	s (light travel time)
15	Orbital eccentricity	dimensionless
16	Barycentric time (TDB scale) of periastron (MJD)	days
17	Longitude of periastron	deg
18	First derivative of longitude of periastron	deg per Julian year
19	Time-dilation and gravitational redshift parameter	s
20	First derivative of orbital period	dimensionless
21	Source of orbital parameters	dimensionless

Performance requirements

Performance requirements will be considered in detail in the full requirements document. This database is not expected to be particularly large or to have particularly demanding access requirements. See the Science Tools Database Requirements Document.

Host environment

Database server system if required (see open issues below). This database should be accessible through a web interface. A user can download this database to his/her server for faster access.

Existing counterparts

Nothing directly applicable. The corresponding files for the EGRET analysis system are plain ASCII files.

Open issues for definition or implementation

There are a couple of ways to organize this database. Two configuration of database are considered feasible at the moment: a single flat table and four structured tables. These options are explained below.

1. a) Single flat table
The database consists of single table with one row for each pulsar entry that contains all the items in the table above. The structure is very simple, but multiple entries for a single pulsar must be allowed, and multiple names for a pulsar (e.g., Crab and PSR0531+21) are not allowed in this scheme.

b) Four structured tables
The database consists of four data tables: a pulsar name table, a spin parameter table, an orbital parameter table, and an originator code table. Each table contains the following items. The pulsar name table is separate from other tables to support multiple names for a pulsar such as the Crab pulsar.

In the tables below, pulsar ID, spin parameter ID, orbital parameter ID, and originator ID are a sequential number uniquely assigned in each table.

Pulsar name table
Primary key: pulsar name
Data item: pulsar ID

Spin parameter table
Primary key: spin parameter ID
Data items: pulsar ID, items 2-12 in the table above. Item 12 is a primary key in the originator table.

Orbital parameter table
Primary key: orbital parameter ID
Data items: pulsar ID, items 13-21 in the table above. Item 12 is a primary key in the originator table.

Originator table
Primary key: originator ID
Data items: name of originator, description of originator, name of contact person, contact information (e.g. postal address, phone number, and e-mail address)

Three options to implement this database have been identified:

1. Single file
The simplest implementation of this database is to have a single file (preferably in ASCII format or FITS). This allows a user to download the entire database onto his/her local machine without any database software installed. However, update, maintenance, and transaction control must be done manually.
Note that either of the two configuration above (flat-table configuration and four-table) can be put in a single file, for example, by using FITS extensions.

2. Conventional database system
This database can fit into any relational database available. This allows easy maintenance, update, and transaction control for the database maintainer. However, in order to download this database to a user's local machine, he/she has to install the same database software as the GLAST team has.

Combined solution: database system at the server, single file at a client
This database can be implemented in a relational database system at the server machine for easy maintenance, while the downloaded database can be in a single file so that a user can access it without any database software. This solves all the issues mentioned above, but it requires the Pulsar Ephemeris Extractor tool (U11) to be capable to handle two different forms of ephemeris database.

The database must be updated by the person(s) that are designated by the GLAST collaboration/operation team. The update procedure is not determined yet.

D5. LAT point source catalog

Date: 13 September 2002 (draft v7 – updated link to requirements document); 25 April 2002 (draft v1); 12/1/02 DLB

The LAT team will produced and maintain a catalog of detected sources in the LAT data. The catalog will be valuable both for further study itself and as a resource for definition of models in U7.

Function

This is the online form of the point source catalog (to be constructed by the LAT team). It is not directly accessed by the analysis tools, but instead receives queries from, and passes data to, the Catalog Access tool (U9).

Contents

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-008, are as follows:

	Contents	Units
1	source name (“telephone number”)	dimensionless
2	RA	deg
3	Dec	deg
4	th68 semimajor, semiminor axis, and position angle	deg
5	th95 semimajor, semiminor axis, and position angle	deg
6	flux (>100 MeV, avg. for the time interval of the catalog)	cm-2 s-1
7	flux uncertainty, 1 sigma (as above)	cm-2 s-1
8	photon spectral index (avg)	dimensionless
9	variability index	dimensionless
10	significance (avg)	dimensionless
11	significance (peak)	dimensionless
12	peak flux (for time interval above?)	cm-2 s-1
13	peak flux uncertainty	cm-2 s-1
14	time of peak flux (wrt reference date)	s
15	interval of time	s
16	flux history	cm-2 s-1
17	flux uncertainty, 1 sigma (as above)	cm-2 s-1
18	start times of flux history entries	s
19	end times of flux history entries	s
20	candidate counterparts	dimensionless
21	degrees of confidence for the counterparts	dimensionless
22	flags (confusion, low latitude,…)	dimensionless

Performance requirements

Performance requirements will be considered in detail in the full requirements document. This database is not expected to be particularly large or to have particularly demanding access requirements. See the Science Tools Database Requirements Document.

Host environment

Database server system

Existing counterparts

Nothing directly applicable. Spatial indexing methods (hierarchical tessellations of the sky) like those developed for the Level 1 database (D1) might be carried over directly to this application, although the size of the database will be so much smaller that spatial indexing probably will not be required for adequate performance.

Open issues for definition or implementation

1. Are there any other quantities that should be included in the catalog entry for each point source? Note that typical sources, i.e., sources not much brighter than the detection limit, will have poorly-determined spectra, and so in general detailed spectral information probably does not belong in the catalog.
2. Should multiple versions of the catalog be defined? In the analysis process, a large catalog of all detections of sources (say, on a daily, weekly, or longer timescale) could be compiled, then that catalog cross-correlated with itself to determine which sources had (probably) been seen more than once. This process would also be useful for searching for transient sources.

D6. Other high-level databases

Date: 10 September 2002 (draft v5, edited by R. Schaefer); 12/1/02 DLB

Function

D6 is a catch-all designation for all other catalogs used by the tools. These catalogs include the GRB and LAT transient source catalogs, which are both largely subsets of the LAT point source catalog. It is also possible that the grand catalogue just compiled for the INTEGRAL mission may be used. This catalog contains information about sources observed by other instruments (e.g., the EGRET point source catalog, and x-ray point source catalogs). D6 will likely be implemented as tables in a conventional database system, such as Oracle and MySQL, or left as ASCII tables to be browsed by U9.

Contents

See outputs for the Catalog Access tool (U9).

The database will contain at least LAT derived burst and transient catalogs. Queries to the database will return the following provisional rows of information based on the Data Products Working Group report.

LAT Transient Catalog:
The output of queries to the LAT Transient Catalog as defined in the report of the Data Products Working Group as LS-007

Column Number	Column Name	Units
1	source name ("telephone number")	dimensionless
2	RA (best position, e.g., from when transient was brightest)	deg
3	Dec	deg
4	th68 semimajor, semiminor axis, and position angle	deg
5	th95 semimajor, semiminor axis, and position angle	deg
6	flux (>100 MeV, avg. for the time interval of the detection)	cm-2 s-1
7	flux uncertainty, 1 sigma (as above)	cm-2 s-1
8	photon spectral index or hardness ratio (avg)	dimensionless
9	variability index	dimensionless
10	significance (avg)	dimensionless
11	significance (peak)	dimensionless
12	peak flux (for time interval above?)	cm-2 s-1
13	peak flux uncertainty	cm-2 s-1

14	time of peak flux (wrt MJDREF)	s
15	interval of time	s
16	flux history	cm-2 s-1
17	flux uncertainty, 1 sigma (as above)	cm-2 s-1
18	start times of flux history entries	s
19	end times of flux history entries	s
20	candidate counterparts	dimensionless
21	degrees of confidence for the counterparts	dimensionless
22	flags (confusion, low latitude,…)	dimensionless
	Note that source name (column 1) should clearly indicate that source is provisional or a transient, e.g., 'GL TR ####+####'	

LAT Burst Catalog:
The output of queries to the LAT Burst Catalog as defined in the report of the Data Products Working Group as LS-009;

Column Number	Column Name	Units
1	GRB name (date encoded name)	e.g., GRByymmdd
2	alert number	running GLAST alert number
3	LAT alert time wrt MJDREF	s
4	GBM alert time wrt MJDREF	s
5	RA	deg
6	Dec	deg
7	th68 semimajor, semiminor axis, and position angle	deg
8	th95 semimajor, semiminor axis, and position angle	deg
9	peak flux > 30 MeV	cm-2 s-1
10	peak flux uncertainty, 1 sigma	cm-2 s-1
11	time of peak flux wrt MJDREF	s
12	energy of max energy photon	GeV
13	energy uncertainty, 1 sigma (above)	GeV
14	time of most energetic photon wrt MJDREF	s
15	duration measure	s
16	duration measure uncertainty, 1 sigma	s
17	duration measure start time wrt MJDREF	s
18	avg photon energy > 30 MeV	GeV
19	Uncertainty (above)	GeV
20	fluence > 30 MeV	cm-2
21	fluence uncertainty, 1 sigma (as above)	cm-2
22	avg photon energy > 100 MeV	GeV

23	Uncertainty (above)	GeV
24	fluence >100 MeV	cm-2
25	fluence uncertainty, 1 sigma (as above)	cm-2
26	rate history > 30 MeV	s-1?
27	rate history uncertainty, 1 sigma (as above)	s-1?
28	rate history bin start time	s
29	rate history bin stop time	s
30	photon spectral index (avg)	dimensionless
31	photon spectral index uncertainty, 1 sigma	dimensionless
32	flags for grb catalog entry	

Performance requirements

This database is not expected to be particularly large or to have particularly demanding access requirements. We expect queries to be answered in seconds rather than minutes.

Other modules required

None.

Host environment

Database server system accessible though the web or alternatively included in a web page served by a public web server. A user can download this database to his/her server for faster access. In that case, he/she must have the database software for this database installed in his/her server machine.

Existing counterparts

Nothing directly applicable. The equivalent information for EGRET was stored in ASCII tables.

Open issues for definition or implementation

1. Shall the INTEGRAL source catalog be included in this database? The INTEGRAL analysis software group is combining information from several source catalogs in different energy bands for use with their science analysis tools. Should we adopt their catalog and catalog updates, or should we go a different route, perhaps making a tool to query other catalogs directly to avoid the need to constantly update the INTEGRAL catalog? Adopting the INTEGRAL catalog will likely require a database management system like MySQL or Oracle.

4 Utilities

Utilities are tools that manipulate data but do not produce an astrophysical result. Some utilities are used to access databases, thereby isolating the databases and the SAE from each other.

In providing the input and outputs for each tool the FITS file type is indicated (e.g., FT1 for event data). The input is separated into "parameters" and "data;" the latter may be input through one of the standard FITS file types.

U1. Event data extractor

Date: 15 September 2002 (draft v9); 11/27/02 DLB (v10)

Function

This is the front end to the Level 1 databases – the gamma-ray summary and the event summary databases. It can be used as a standalone tool to produce FITS files for export outside of the analysis system, or it can be run by standard high-level analysis tools. The latter will generally run on the data requestor's computer while the Data Extractor will run on the SSC or LAT team-based server computers hosting the databases.

The Data Extractor utility constructs queries to the Level 1 database (D1) and writes FITS files containing the retrieved events.

Inputs (for gamma rays)

Parameters:

- Time range
- Region of sky (specified as center and radius or coordinate ranges)
- Energy range
- Zenith angle cuts (energy dependent, and possibly also on inclination and azimuth or even plane of conversion[1])
- Inclination angle range
- Azimuth range
- Gamma-ray type (assigned in Level 1 processing based on sets of background rejection/PSF enhancement cuts that have corresponding IRFs tabulated in D3, CALDB)
- Solar system object flag (for indicating whether a suitable radius around solar system objects – the moon and sun in particular - should be excluded from the selection)
- The coordinate system for the search should be selectable, among celestial, Galactic, instrument-centered, earth-centered (and possibly sun and moon-centered).

Note that the selection criteria specified here are identical to the criteria supplied to the Exposure Calculator to calculate the corresponding exposure.

---

[1] Cuts on zenith angle will be required to eliminate albedo gamma rays produced in the upper atmosphere. Crude cuts on zenith angle will be applied onboard (to limit the impact of albedo gamma rays on the average data rate). However, more detailed cuts will have to be applied in ground processing. Owing to the strong dependence of the PSF size on energy, to a lesser extent on inclination and azimuth, and other parameters like plane of conversion), more conservative cuts must be applied at lower energies (and larger inclinations) for the same rejection efficiency. These same cuts must also be incorporated into the exposure calculation. These cuts may be difficult to specify as selection criteria for extracting events from the databases, a more practical approach may be to assign an 'albedo-ness' value to each event during Level 1 processing. The albedo value (basically yes or no) could be specified as an input to the Data Extractor. If this implementation is adopted, care must be taken to ensure that the Exposure Calculator has the same algorithm as the Level 1 processing pipeline for the albedo cuts.

Databases required

Level 1 event summary D1

Outputs

- FT1—the selected events. The output must include as part of the header the complete specification of selection criteria.

Performance requirements

The requirements for the Data Extractor are distinct from the requirements for the Level 1 database itself. With the query processing being handled by the data server, the principal processing work done by the Data Extractor should amount to reformatting the data that have been extracted from the database. Requiring only that the overhead for this work be much less (<10%) than the time required by the database system for a typical query is sufficient.

Other modules required

GUI front end (if we define such a thing as a general utility for providing user interfaces to server-run tools)

Host environment

Database server system (with user authentication for proprietary data access

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

1. Should all of the coordinate systems listed above be available, and which should be the primary system? (Database performance is likely to be significantly greater for the primary system if data are sorted by the corresponding spatial index before they are ingested.)

2. As currently conceived, the event summary database contains summary information (output of reconstruction and classification) for all telemetered events and the gamma-ray summary database contains only those events that are classified as gamma rays by at least one of the standard sets of classification cuts. Conceptually, it would be cleaner to keep all events together in one database, but at least two good reasons indicate that a separate gamma-ray database is desirable for enhanced performance: gamma rays will represent only ~10% of the total events, and they will be most usefully accessed by region of the sky (as opposed to arrival direction in instrument coordinates for cosmic rays).

3. Pulsar phase (involving specification of a pulsar with known timing information) would also be a useful selection criterion, but implementation is TBD. A more practical approach may be to make a standalone tool for subsetting selections and scaling exposures.

4. Selection criteria related to, e.g., geomagnetic cutoff might also be useful if it turns out that residual cosmic rays are a problem. What geomagnetic parameters will be the most useful in this regard?

5. The number of events returned by a search could easily be in the millions. Is this a problem in terms of staging data for transfer and transferring it over the network?

U2. User-Level Data Extraction Tool

Date: 16 Sep 2002 (draft v2); 11/27/02 DLB (v3)

Function

This is an interactive environment for applying additional selection criteria (cuts) to LAT event files after they have been downloaded from the central server via U1. The user can specify cuts on any quantity that is contained in the event file, such as position, energy, inclination, time, event quality, and telescope mode. After a given series of cuts, the user will be able to plot the appropriate quantities (e.g., histograms of events as a function of position or energy) to examine the cumulative effect of the selections. At any point, the user can remove or add a cut, and specify parameters such as binning in time, energy, or position. The final output will be a filtered event file that can be fed to other analysis tools. The cuts include energy-dependent (and inclination angle dependent) cuts based on angular distance from a given position; these are used to enhance the pulsed fraction of the emission in periodicity searches of sources embedded in bright diffuse emission.

Inputs

Data:

- FT1—event file(s) produced by U1, with specifications of the cuts in the header

Outputs

- FT2—filtered event files. The FITS headers of these files will contain a complete history of the selection criteria that have been applied, including those specified during the U1 extraction.

Performance requirements

Performance will be limited by the hardware (disk read and write access times) on the client's computer. The software should be able to complete its job within a factor of four of the combined disk read and write times (plus some relatively small system overhead) to read the U1 supplied FITS file and write a new one. Testing of reading and writing of FITS files using cfitsio indicates that this should be easily achievable.

Other modules required

Plot display (UI2) and GUI front end (UI5)

Host Environment

Client computer.

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

None

User Interface

1. A command line interface shall be the basic means of interaction for the user.
2. Scripting capability shall be provided to allow for automation of repetitive tasks and complex sequences of cuts.
3. A GUI shall also be provided as an alternative interface for running this tool. Among its features, the GUI will have pull-down menus describing the event attributes that can be selected.
4. Interactive plotting and image display shall be provided. This will allow, for example, mouse-driven rescaling of the plotting region.

U3. Pointing/livetime history extractor

Date: 1 May 2002 (draft v1); 11/27/02 DLB (v2)

Function

This is the front end to the Pointing, livetime, and mode history database D2. Most commonly, this will be run by the Exposure Calculator U4 to extract the relevant portion of the pointing/livetime history for a particular exposure calculation. It could also be used as a standalone tool to produce FITS files for export outside of the analysis system, but the information is likely to be useful only for exposure calculations.

Inputs

Parameters:

- Time range

Databases required

Pointing, livetime, and mode history D2

Outputs

- FT2—pointing/livetime history database for the selected time interval.

Performance requirements

With the query processing (minimal as it is) being handled by the database server (for D2), the principal processing work done by the Pointing/livetime extractor should amount to reformatting the data that have been extracted from the database. Requiring only that the overhead for this work be much less (<10%) of the time required by the Exposure calculator U4) for a typical query is probably sufficient.

Other modules required

GUI front end (if we define such a thing as a general utility for providing user interfaces to server-run tools)

Host environment

Database server system

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

1. Should the operating mode of the LAT be an input? The idea is to incorporate flexibility for different operating modes (e.g., in the event of a tower failure) that correspond to completely different sets of instrument response functions.

2. The high-level analysis in general, and the exposure calculation in particular, may be based on a tessellation scheme for subdividing the sky. (The tessellation also provides a spatial index for the Level 1 database D1.) Exposure calculations in particular will be much faster for having precalculated values of the angles between each cell and every other cell in the tessellation. Should the pointing/livetime history extractor include the cell values for the relevant directions (z-axis, y-axis, zenith) in the portions of the timeline that it extracts? The alternative would be to have these directions translated into cell index values by the Exposure calculating tool U4.

3. Pre-extracted pointing/livetime histories for some standard time intervals (e.g., monthly or yearly), or whatever intervals appear to be commonly used for analyses could be made available.

U4. Exposure calculator

Date: 7 May 2002, v3 (22 Apr 2002 v1; 1 May 2002 in v2 clarified the function of the tool, its inputs, and a possible implementation; 7 May 2002 in v3 further clarified inputs & outputs); 11/27/02 DLB

Function

This module calculates the exposure for a given sky region over a given time range. The exposure can be provided at different levels of detail, i.e., as a function of different detector parameters. Exposures are necessary inputs to most high-level analysis tools and are also required for Observation simulator O2.

Inputs

Parameters:

- Desired detail of exposure map; user may want the total exposure or the exposure as function of detector parameters such as inclination angle, so a set of parameters to be 'marginalized' can be specified (see Open Issues below)

Data:

- FT1—event list from U1 or U2. Header includes specification of time range, energy range, inclination angle range, zenith angle cuts (and their energy dependence), photon type or quality, possibly even cuts on geomagnetic parameters and cuts on the directions of the sun and moon, etc.
- FT2—Pointing/livetime history extracted by U3 from Pointing/livetime database D2 or generated by pointing/livetime simulator O1

Databases required

D3 - CALDB for the effective area component of the IRFs and for possible application to the zenith angle cuts (but see Open Issues below).

Outputs

- FT7—the exposure

The exposure array or matrix (i.e., a map with multiple dimensions) will be stored at the desired level of detail. The specification of parameter ranges on input and the versions of the instrument response functions used must be recorded with the output, e.g., as part of the header.

The exposure itself could be thought of as a large table of values (cm^2 s) as a function of spatial cell in the region of interest, and a grid of inclination, azimuth, energy, event class, and any detector parameters that are included in the definitions of the instrument response functions (perhaps like plane of conversion). For some of the dimensions, parameterization of the variation of the exposure may be possible. In this case, the exposure array would contain coefficients of the parameterization, and the table would be a lot smaller.

Performance Requirements

TBD, but at a minimum time required for execution should be comparable to the Data Extractor's retrieval time for the corresponding gamma-ray data.

Other Modules Required

Pointing/livetime history extractor U3

Host Environment

Run on central server or client computer.

Existing Counterparts

None. The calculation of exposures for EGRET (by the program INTMAP) would require extensive revision to accommodate the exposure calculations for the LAT, which are more complicated in some respects (e.g., number of parameters in the instrument response functions) and simpler in others (e.g., only one basic operating mode for the LAT vs. dozens of combinations of EGRET's trigger telescopes).

Open Issues for Definition or Implementation

1. A user may want a time history of the exposure. Should this be implemented as an option within the exposure calculator, or should the calculator be run multiple times for different (adjacent) ranges of time?

2. Pointing/livetime information also can be used to derive the temporal window function (i.e., visibility profile) for a specified direction and time range. Should the extraction of window functions be an option in the Exposure calculator tool?

3. Spatial access to the gamma-ray data will likely be facilitated using a tessellation of the sky. Each tile in the tessellation is assigned an index. Each tile will be small enough (less than degree sized) that the exposure does not vary significantly across a tile. The analysis tools, e.g., for point-source detection may be implemented to use the same indexing scheme for the exposure, in which case the standard exposure matrix would not be a map but rather a tabulation (of multidimensional quantities).

4. The exposure 'map' may have so many dimensions to be unwieldy to tabulate for any particular analysis. The calculation of exposure can be factored somewhat into tabulation of total livetime for each spatial cell as a function of inclination, azimuth, zenith angle, and instrument mode (different modes corresponding to different sets of instrument response functions) which would be the output of this tool. (If it is determined that we will need to include geomagnetic parameters in the cuts for the analysis, then the tabulation must also include them. The tabulation would be even more unwieldy.) Then any other tool desiring exposures could calculate them from this tabulation plus a specification of the response functions (energy and angle dependence of the effective area) along with the prescription for the zenith angle cutoff. This approach adds flexibility, too, because for the same gamma-ray data set, different response functions

might want to be considered, e.g., for gamma rays that convert in the front vs. the back of the tracker. The same tabulation of livetime distribution can be applied to each.

5. Exposure calculations for time intervals short relative to the ~30–60 s interval between entries in the Pointing/livetime database (D2) will require special treatment. The orientation and position of the LAT will not vary enough over such short time intervals to make a significant difference in the exposure calculation, but the accumulated livetime can be determined only by examining the livetime counter values recorded with each event in that time interval. To a good approximation, these accumulated livetimes can be used to scale the overall exposure for the time interval between entries in the Pointing/livetime database. An exposure cannot be associated with a particular gamma ray; the shortest time intervals for which exposure can be calculated correspond to the shortest intervals between two events that are telemetered from the LAT. A special tool for calculating exposures on short time intervals, or perhaps for constructing profiles of livetime accumulation during short intervals, might need to be defined.

U5. Interstellar emission model

Date: 26 April 2002 (draft v1)

Function

This is the model of the interstellar gamma-ray emission of the Milky Way used both for analysis of gamma-ray data and generation of simulated data. For both applications, the distribution of interstellar intensity on the sky and in energy is one component of a more detailed model of the gamma-ray intensity that includes point sources and small extended sources. The interface to the interstellar emission model will be the source modelling module [which must have an option to write out a model].

The model of interstellar emission will be defined with TBD adjustable parameters, most likely related to the spatial and energy distributions of cosmic rays. In principle, the optimum values for these parameters will be determined from global analysis of LAT data and for subsequent studies these parameters could be assigned fixed values. However, owing to the need (at least initially) for adjustability of the parameters, the Interstellar emission model cannot be planned as a static database as far as the analysis environment is concerned.

Keep in mind that the requirements specified here apply only to the module that provides the interstellar intensities to the source modelling module. The actual development of the interstellar emission model (via models of the interstellar matter and radiation and cosmic rays) is not the subject of these requirements.

Inputs

The inputs are the region of the sky and energy range, along with the values of the TBD adjustable parameters. Depending on the implementation of the source modelling module, the coordinate system, gridding, and energy gridding of the model might also be inputs.

Databases required

NA

Outputs

The provisional outputs, defined in the report of the Data Products Working Group as LS-010, are shown in the table below. The DPWG approach assumes that the coordinate grid on the sky will be the same as used for generating the exposure and indexing the gamma rays for spatial access. For completeness, the output may also include the celestial coordinates of the spatial grid and the specific energies for the model. The output should also include specification of the version of the underlying model of the interstellar emission and the values of the adjustable parameters.

The DWPG assumed that in addition to the intensity, the Interstellar emission model would also provide a specification of the uncertainty in the model. This has not been explored in terms of the overall analysis system. Whether the uncertainty represents a statistical uncertainty or a systematic uncertainty is not yet specified, nor is how these uncertainties would be quantified. Neither has the possible incorporation of this information in model fitting has not been investigated.

Keep in mind that the output of this module will not have been convolved with any instrument response functions. Also, the intensities are differential. (For EGRET analysis, integral intensities were calculated for standard energy ranges.) This allows for more straightforward interpolation in energy and will also facilitate spectral analysis within the general Likelihood analysis tool A1. (If fluxes are modelled on integral energy ranges, as they were for EGRET, then spectra must be derived after the fact by ‘forward folding’ models to find agreement with the integral fluxes found from the likelihood analysis.)

	Column Name	Units
1	pixel number	dimensionless
2	intensity spectrum	photon cm-2 s-1 sr-1 GeV-1
3	intensity uncertainty (1-sigma)	photon cm-2 s-1 sr-1 GeV-1

Performance requirements

Performance requirements will be considered in detail in the full requirements document. No particularly demanding requirements are apparent.

Other modules required

None

Host environment

Database server system

Existing counterparts

Nothing directly applicable. For EGRET analyses, the models were entirely precomputed

Open issues for definition or implementation

1. In terms of the Likelihood analysis tool A1, the model may need to be recomputed many times during a given analysis with different values of the adjustable parameters. If the model depends non-linearly on the parameters, there is not a lot that we can do to speed up the process. However, if the dependence on any of the parameters is linear, and if performance is a concern, then separate model components for each of the linear terms should be returned as output.

2. Other open issues, related to the assumptions made by the DPWG in defining the output of the Interstellar emission module, are included in the Output section above.

U6. Map generator (DS9?)

Date: 7 Sept 2002 (23 April 2002 draft v1 updated with Y. Ikebe's comments about intensity maps); 11/27/02 DLB

Function

This module constructs binned maps of gamma rays ('counts'), exposure, and intensity from the output of the Event data extractor (U1) and the Exposure Calculator (U4) or from the interstellar emission model. Calculations of exposure maps will likely involve interpolation of exposures calculated for the standard tessellation grid. Exposure maps for finite energy ranges will have to calculated for an assumed source spectrum, to permit the correct relative weighting of the energy-dependent effective areas. Intensity maps for finite energy ranges can be calculated without an assumed spectral form if they are calculated as the sum of counts/exposure ratio maps for ranges of energy that are narrow enough that the effective area does not change significantly. U6 shall be able to calculate intensity maps in this way.

Inputs

Parameters:

- Coordinate gridding of the maps and coordinate projection.
- Parameters to be 'marginalized' (integrated over)
- Gridding of the remaining parameters (e.g., energy or inclination angle)
- Smoothing factors (e.g., for spatial convolution with a gaussian)

Data:

- FT1—event list produced by the event data extractors U1 or U2
- FT7—Exposure matrix produced by the Exposure calculator U4

Databases required

Interstellar emission model (U5)

Outputs

- Counts and/or exposure maps (perhaps multidimensional). The intensity map, i.e., the ratio of counts and exposure, can be calculated as well. All output must include (in a header) the specifications of selection criteria used (for the input gamma-ray and exposure datasets) as well as of the parameters marginalized as part of the map generation.

Performance Requirements

TBD. The map generation is not likely to be computationally intensive.

Other Modules Required

None.

Host Environment:

Run on central server or client computer.

Existing Counterparts

None. The EGRET analogs, MAPGEN (counts maps) and INTMAP (exposure and intensity maps), are based on single viewing periods,

Open Issues for Definition or Implementation

1. What coordinate systems should be supported? Systems other than in the photon dataset?

2. What coordinate projections should be supported?

3. Will any analysis tools require as input any of the maps generated here as input? Or will the maps primarily be for visualizations?

U7. Source model definition tool

Date: 18 Aug 2002 (draft v1.2)

Function

This tool enables the user to define a source model for simulations and model fitting. The output of U7 comprises a collection of parameter values and identifications of the referenced model components and shall provide information regarding the gamma-ray source model to the Observation Simulator (O2) and the Likelihood Analysis Tool (A1).

Model Specifications

Source models shall consist of the following components and specifications:
1. Diffuse Emission Model Parameters
(a) Galactic Diffuse
    i. Parameter values
    ii. Descriptor(s) of the interstellar emission model (version and/or filename)
(b) Extragalactic Diffuse Parameters (flux and spectral index)
2. Point Sources
(a) Source name(s)
(b) Location(s)
(c) Spectral Properties
    i. parametric representation
    ii. template file (spectrum file)
(d) Variability Properties
    i. parametric representation
    ii. template file (light curve file)
3. Extended Sources
(a) Source name(s)
(b) Spatial Properties
    i. parametric representation
    - location of reference point
    - shape, size, orientation angle
    ii. template file (map of intensities)
(c) Spectral Properties
    i. parametric representation
    ii. template file

Standard Model Components

A set of standard model components, such as power-laws for photon spectra, wavelet basis functions for light curves, or annuli for shapes of extended sources, shall be made available. The spectral and spatial components will be either additive or multiplicative in nature. For example, to model a blazar that has two power-law components, each one resulting from a different emission region along the jet axis, and absorption from the intervening intergalactic infrared radiation, the two power-laws are additive components and the absorption is multiplicative. Mathematically, this model is expressed as

$$dN/dtdAdE = e^{-\tau(E,z)} (N_1 (E/E_1)^{-\alpha_1} + N_2 (E/E_2)^{-\alpha_2}). \qquad (1)$$

The optical depth τ is a function of photon energy and the source redshift z; and N_i, E_i, and α_i are the normalization, reference energy, and photon spectral index of the ith power-law component. Temporal components can be both additive and multiplicative. For example, a source may have separate periodic and flaring components whose fluxes are to be added together. Such a source may be represented as a single source with two additive temporal components or as two distinct sources with the same location on the sky. The latter representation would allow the user to define different spectral properties for the two temporal components whereas the former would not. Alternatively, in order to model the gamma-ray analog of the bursting X-ray pulsar, a flare component and periodic component are combined multiplicatively.

Importing Parametric Models

The capability to import user-defined parametric models must be provided. This should be similar to the "import" capability of the isis spectral analysis program (http://space.mit.edu/CXC/ISIS/) that allows the Xspec models to be used for spectral fitting in addition to the native ISIS models. In ISIS, one simply executes the command

```
ISIS> import ("xspec");
```

and all of the Xspec models become available. See Section 8 of the ISIS manual, ftp:// space.mit.edu/pub/CXC/ISIS/manual.pdf.

This will require the I/O specifications of the spectral model routines to be carefully defined and provided to the users for their implementation.

Valid Parameter Ranges

Nominal valid ranges and absolute valid ranges must be defined in the software for all model parameters. The nominal ranges may be adjusted by the client when they specify the parameter values, but the nominal valid ranges must always lie within the absolute ranges. The absolute ranges, such as requiring optical depths to be nonnegative (i.e., no gamma-ray masers), cannot be changed. Imported models must also have predefined nominal and absolute ranges. In addition, the user may specify an overall time interval, energy range, and/or region of the sky for which the model is valid.

Default Configuration

The user shall be able to define default parameter values and nominal ranges for any model component. These default values shall be stored in a configuration file that is read by U7 on start up. The user can modify these default values by editing the configuration file directly or via the GUI. Default specifications of the interstellar emission model and various template files may also be specified by the user and shall be stored in the configuration file. If a configuration file does not exist, a configuration file can be created, at the user's option, by U7. The configuration file can be specified by an environment variable. If such an environment variable does not exist, then U7 will look for the configuration file in the current working directory.

Models Defined by Template Files

Model components defined by template files shall also be either additive or multiplicative. Spectral or temporal template files may either be in the form of a standard OGIP FITS file such as .lc files, or they may consist of columns of ASCII data. Spatial template files must be FITS images. Template files will also have parameters associated with them. At a minimum, these parameters will specify the following information:

Spectra

E0**:** a fiducial energy

N0**:** the normalization at E0

Light Curves

t0**:** absolute starting time

tscale**:** scaling factor

pflag**:** flag to indicate periodicity

Extended Shape Maps

long**:** longitude of reference point in chosen coordinate system

lat**:** latitude of reference point

Model Specification

The model for individual sources will typically be constructed interactively using the GUI. The GUI menus will list all of the available models. In order to combine various model components for a given source, a syntax similar to that used by the X-ray spectral fitting program Xspec will be used (see http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/manual/manual.html and http://heasarc.gsfc.nasa.gov/webspec/webspec advanced.html). The following is an example of the use of this syntax:

```
PtSrc(1).name = 'transient_1';
PtSrc(1).ra = 205.;
PtSrc(1).dec = -30.;
PtSrc(1).spec = IGM_absorb(tau0=1, z=0.5)
*(powerlaw(N0=1e-4, E0=1., Gam=-2.1, Gam.min=-3., Gam.max=-1.))
PtSrc(1).time = fred(t0 = 1D3, Trise=1e2, Tdecay=1e4);
ExtSrc(1).name = 'SNR_1';
ExtSrc(1).shape = annulus(ra=200., dec=-32., rinner=0.5, router=0.75);
ExtSrc(1).spec = powerlaw(N0=4e-5, E0=1., Gam=2.);
PtSrc(2).name = 'PSR_1';
PtSrc(2).ra = 200.;
PtSrc(2).dec = -32.;
PtSrc(2).spec = template(file='psr_spec', NO=2e-5, E0 = 1.);
PtSrc(2).time = template(file='psr.lc', t0=1.2345D3, tscale=1.,
pflag=1);
```

For extended sources, the normalization N0 refers to the flux integrated over the entire annulus. For varying sources, it refers to the peak flux. Nominal ranges can be provided at the user's option. The names of the nominal range variables are given by appending

.min and .max to the parameter value name. Parameter values can also be specified to be held fixed during the likelihood analysis. For example, to hold the spectral index of the extended source fixed in the above example, one adds Gam.fixed = 1 to the parameter list; to specify explicitly that it is not fixed, one enters Gam.fixed=0. As with all parameters, the default values for whether a parameter is fixed or not can be specified in the configuration file by the user.

Model Verification

The Source Model Definition Tool shall verify that the model parameters are within their valid nominal ranges, that the nominal ranges are within the absolute ranges, and that the requested template files are available and contain valid data. If a nominal range value lies beyond its corresponding absolute range value, it will be set at the absolute range value and a warning will be issued. U7 shall also verify that the various components are used in their proper contexts; in particular, it will ensure that uniquely additive components are not used in a multiplicative context and vice versa.

User Input

- User input shall be performed through a GUI.
  Or
- FT8—model files can be edited directly by the user and those files can be read by U7 for verification.

Model Visualization and Interactive Manipulation

The GUI shall provide a means of displaying actual gamma-ray data for the analysis region of interest. These data will consist of counts or intensity maps from the EGRET archive or from the LAT data itself. The GUI will have the ability to indicate the locations of the various model components by overlaying appropriate representations of each component on the counts or intensity maps. For point sources, plotting symbols will be used; for extended source shapes, the outline of the shapes will be plotted; and for the spatial templates of extended sources, contour plots will be overlaid. The location, and for shapes, the size and orientation, of each component will be adjustable interactively using the mouse, and the corresponding parameters will be updated automatically. A window or frame shall be provided by the GUI that lists the model parameters, and those parameters shall also be accessible to the user for direct editing in the window or frame. In addition, a counts or intensity map of the model, convolved through the instrument response, may also be displayed.

Required Ancillary Data

The Source Model Definition Tool shall have access to the following data.

- Output from the Interstellar Emission Model (A5)
- LAT point source catalog (D5)
- Other high-level databases (D6)
- CALDB (D3) for constructing model intensity maps that have been convolved through the instrument response

Outputs

- FT8—The principal output of this tool shall be an XML file containing a complete definition of the source model. This shall include the model components and parameters that are explicitly specified by the user as well as the default parameters that are either provided by U7 itself or from the user's configuration file. In this way, the output model file will contain a complete description of the source model that is independent of other files, except the template files. The overall ranges in time, energy, and region of validity for the model as optionally supplied by the user shall also be included in the output XML file. In addition, model counts or intensity maps may be generated at the user's option as FITS files, and any of the displayed images, such as maps with model overlays may be created as images in appropriate formats (JPEG, GIF, PostScript, etc.).

Host Environment

Client computer.

Open Issues

1. We need a way of distinguishing the parameters for different components of a single source that are described by the same model. The two additive power laws in equation (1) is an example.

U8. IRF visualization

Date: 6 September 2002 (draft v1)

Function

This utility extracts response functions from D3 for visualization. It is not conceived as a general front end for D3, i.e., a utility that stands between it and the analysis tools that need instrument response functions. (The CALDB implementation mandated for D3 makes an interface utility per se unnecessary for post-mission maintenance of the software by the HEASARC. A library of routines for directly accessing the response functions in D3 will undoubtedly be used by the analysis tools, and by U8.) Instead, it is a tool for inspection of the response functions.

Inputs

Specification of the response function (effective area, energy resolution, point-spread function)
Specification of the relevant parameters (e.g., event class, energy range, angle ranges)

Databases required

Instrument Response Functions D3.

Outputs

Plot or tabulation in a file of the requested response function.

Performance requirements

No special requirements.

Other modules required

None

Host environment

Run on client computer. (Regardless of whether D3 is installed locally, standard methods exist for remote access to CALDB.)

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

None

U9. Catalog access

Date: 13 September 2002 (draft v8 updated required databases);12 September 2002 (draft v5 –update by R. Schaefer)

Function

This module is a catch-all front end for the catalogs of astronomical sources to be made available as part of the LAT analysis environment. The kinds of catalogs will include the LAT burst and transient source catalogs, the EGRET point source catalog, flat spectrum radio sources, radio pulsars, SNRs, OB associations, blazars, etc. The idea would be to make available the catalogs that would be of interest for gamma-ray studies. It is likely that the catalogs would be stored on a database server, although in principle (especially for any large or dynamic catalogs) access requests for remote servers could be managed. This module should have both an interactive and an API. The search criteria for the catalogs will at a minimum include region of the sky, but could also be specific to each particular catalog.

Inputs

Source name or 2-d location
Region of the sky
Catalogs of interest
Subselection parameters for the individual catalogs; this may include specifications of what information from each catalog is to be returned.

Databases required

D5, D4, and the collection of useful catalogs represented by D6, and their associated access methods.

Outputs

Tables of sources extracted from the catalog(s). (Draft FITS headers for the LAT catalogs are available in the report of the DPWG.) The output must include as part of the header the complete specification of selection criteria.

Performance Requirements

No obvious requirements; queries are not likely to take long to process, i.e., responses should be almost at interactive speeds for regions of the sky of 10’s of square degrees.

Other Modules Required

None.

Host Environment:

Run on central server or client computer. The catalogs themselves will be stored on the server if they are not already served from an external remote site.

Existing Counterparts

Should be investigated carefully. NED?

Open Issues for Definition or Implementation

1. The output should in principle be readable by the model-definition module. This is probably not difficult, just a matter of using consistent naming schemes for the columns so the model-definition module can recognize coordinates and fluxes. (The model-definition module is currently defined under O2, the observation simulator, but is also relevant for the likelihood analysis tool A1.)

U10. Photon arrival time converter

Date: 9 September 2002 (v2, M. Hirayama)

Function

This tool converts photon arrival times at the satellite in a photon list to those at the geocenter, and those at the solar system barycenter according to the satellite location in reference to the geocenter and the Earth location in reference to the solar system barycenter. It also converts barycentric arrival times to binary-demodulated arrival times for binary pulsars. For user's convenience, it performs the time conversion for a single photon arrival time (at the satellite or at the geocenter) given by a user.

Inputs

Parameters:

- RA and Dec of the pulsar to be used for arrival time conversions

Data:

- FT1—event list from U1 or U2
  Or
- A single photon arrival time at the satellite or at the geocenter

Databases required

Earth's ephemeris (e.g. earth.dat in FTOOLS)
(Note: satellite location needed for geocentric correction is available in an event file)

Outputs

Given a photon list as input:

- FT1—event list with the addition of geocentric arrival times , barycentric arrival times (for all pulsars), and a binary-demodulated arrival times (for binary pulsars only). If photon data are from an epoch outside the validity of the satellite's orbital parameters or of the Earth's ephemeris, a warning message should be displayed.

Given a single photon arrival time:

- The arrival times at the geocenter and at the solar system barycenter.

Performance requirements

This should be a relatively simple tool that should run fast.

Other modules required

None

Host environment

Client's server

Existing counterparts

This is a mission-specific tool, but similar tools are available for other satellite missions such as ASCA (TIMECONV of FTOOLS) and XMM (FXBARY of FTOOLS).

Open issues for definition or implementation

1. Should the arrival time corrections for binary pulsars include the relativistic Shapiro delay?

U11. Pulsar ephemeris extractor

Date: 9 September 2002 (draft v2, M. Hirayama)

Function

This is the front end to the Pulsar Ephemerides database (D4). It constructs queries to the Pulsar Ephemerides database (D4) to search for pulsar parameters of a known pulsar at any given time during the LAT mission, and outputs a file in a certain format (e.g. a FITS file, an ASCII file) containing pulsar parameters that are valid at the time. The file will be used by the Pulsar Analysis tools (U10, U12, A3, and A4) or browsed by a user. It also calculates several quantities derived from selected pulsar ephemerides, such as pulse frequency at a time of interest, for the Pulsar Analysis tools (U10, U12, A3, and A4) and for user's reference.

Inputs

The input parameters to this utility are following. The first two items are selection criteria for ephemeris search; if either of them is omitted, all the entries for the omitted criteria will be selected. The third item is for calculations of derived quantities such as pulse frequencies at a time of interest. If the third item is omitted, no calculations will be performed.

Pulsar name if known, or region of sky to search (specified as center and radius or coordinate ranges)

Time of observation, or the time for ephemeris search

Time of interest, or the time for pulse frequency calculation (if different from the time of observation)

The coordinate system for the sky search should be selectable, among equatorial and Galactic. The time system for the time of observation and the time of interest should be selectable, among Mission Elapsed Time, UTC (Coordinated Universal Time), TAI (International Atomic Time), TDT (Terrestrial Dynamical Time), or TDB (Barycentric Dynamical Time).

Databases required

Pulsar Ephemerides database D4.

Outputs

A table containing the selected sets of pulsar parameters. Note that there may be multiple sets of valid parameters in some cases. The output must include as part of the header the complete specification of selection criteria. In addition, the header should include the following supplementary information for user's future reference.

Time of observation, or the time for ephemeris

Time of interest, or the time for pulse frequency

The table part consists of one or more sets of the following items.

Column	Contents	Units
1	Ephemeris ID	dimensionless
2-22	Items 1-21 in Pulsar Ephemerides database (D4)	
23	Pulse frequency at the time of interest	Hz
24	First derivative of pulse frequency at the time of interest	Hz^2
25	Pulse phase at the time of interest	dimensionless
26	Apparent pulse frequency at the time of interest (*)	Hz
27	Apparent first derivative of pulse frequency at the time of interest (*)	Hz^2
28	Apparent second derivative of pulse frequency at the time of interest (*)	Hz^3

(*) Taken into account of Doppler effect due to pulsar's binary motion. 'Apparent' means apparent to a fictitious observer at the solar system barycenter.

Performance requirements

Performance requirements will be considered in detail in the full requirements document. The Pulsar Ephemerides database (D4) is not expected to be particularly large or to have particularly demanding access requirements.

Other modules required

GUI front end

Host environment

Run on central server or client computer.

Existing counterparts

Nothing directly applicable

Open issues for definition or implementation

Depending on implementation of the Pulsar Ephemeris database (D4), this tool must be capable to access the database in any format(s) defined. For example, if the database (D4) is supplied in two forms, one in a conventional database system and the other in a single file, then this tool must automatically detect a form of the database, and access it accordingly. Actual implementation will be determined based on the choice of database implementation.

Preferably, various output formats should be available for user's choice: extracted ephemeris can be stored in a FITS file, stored in an ASCII file, displayed on screen, and so on. This feature is important especially for use with existing analysis tools other than the tools described here, such as XRONOS in FTOOLS, because outputs of this tool should be available for such external tools, too.

U12. Pulsar phase assignment

Date: 4 September 2002 (v3; edited by M. Hirayama; formerly A3); 11/27/02 DLB

Function

This tool assigns pulse phases to photons according to an ephemeris. The ephemeris may be for a known pulsar or may be input (e.g., the result of a period search). It also assigns orbital phases to photons for binary pulsars.

Inputs

Parameters:

- Pulsar ephemeris for a known pulsar (from D4, extracted by U11), or a result of pulsar search for an unknown pulsar (from A3 or A4).
- Choice of photon arrival times (at the satellite, at the geocenter, at the solar system barycenter, or binary-demodulated)

Data:

- FT1—event list, processed by U10

Databases required

Pulsar Ephemerides (from D4, extracted by U11) when tool is applied to a known pulsar

Outputs

- FT1--event list with the addition of an assigned pulse phase (for all pulsars) and an assigned orbital phase (for binary pulsars only). If photon data are from an epoch outside the ephemeris' validity, a warning message should be displayed.

Performance requirements

This should be a relatively simple tool that should run fast.

Other modules required

None

Host environment

Client's server

Existing counterparts

This is a standard piece of software that must exist in many forms.

Open issues for definition or implementation

1. Should the phase calculations for binary pulsars include the relativistic Shapiro delay?

U13. GRB visualization

Date: 9 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tool plots GRB lightcurves and spectra using LAT, GBM and/or other missions' data. The light curve plots include:
Counts binned in time and energy vs. time
Counts binned in time and energy normalized by the bin livetime vs. time
Time histories of spectral parameters from fits to the spectra
Spectra include:
Raw counts accumulated over specified time range vs. energy bin
Counts accumulated over specified time range normalized by the livetime and the energy bin width vs. energy bin
#2 compared to model count spectrum
Data and model "photon" spectra

Here "counts" are the detected photons without corrections for detector efficiency.

Model count spectra are calculated by folding the spectral model (e.g., from A8) through the relevant detector response (e.g., the ARF and RMF files). The data "photon" spectrum is a scaling of the actual count spectrum by a model-dependent estimate of the detector efficiency.

Plots should be presentable as separate plots or stacked on a single plot. Scripting should allow recreating a previously developed plot.

Inputs

Data:

- FT1—LAT event list (from D1, extracted by U1, perhaps modified by U2; a special photon cut for bright, localized sources may be used) in the standard photon FITS format (see Outputs section of D1)
  Or
- GBM count data in the standard FITS format
  Or
- FT3—LAT binned data (from A5) in XSPEC PHA format
- FT4—ARF for LAT binned data
- FT5—RMF for LAT binned data
  Or
- FT3—GBM binned data (from either A5 or continuous data accumulated by the GBM) in XSPEC PHA format
- FT4—ARF for GBM binned data

- FT5—RMF for GBM binned data
  Or
- FT3—Data from other missions (e.g., Swift, AGILE) in XSPEC PHA format
- FT4—ARF for other missions
- FT5—RMF for other missions

- U13—spectral fits from A8 or A9

Databases required

None

Outputs

- Plots on the user's screen, stored in a file, or sent to a printer.

Performance requirements

Plotting should be fast. The interface should allow easy manipulation of the plots.

Other modules required

None

Host environment

Client computer

Existing counterparts

Many packages (e.g., SOAR , RMfit) plot burst lightcurves and spectra

Open issues for definition or implementation

How much flexibility should the visualization tool provide the user? Should users be able to change the plotted ranges or the axis legends? Should they choose logarithmic or linear scales?

U14. LAT GRB DRM generator for binned count spectra

Date: 9 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tool will calculate the detector response matrix (DRM) for each count spectrum in the series of spectra that span a burst. Since the same matrix can be used for time bins accumulated over periods during which the LAT's inclination angle to the burst changed very little, the DRM may not be calculated for each spectrum.

Burst analysis can be performed on the counts accumulated over a region around the burst; spatial information is not required for the comparison of the data and the model. Thus a simplified instrument response function (IRF) will suffice. The DRM is a matrix relating the incident photon flux, binned in energy, to the observed counts from the region around the burst, binned in apparent energy. The DRM is related to the LAT IRF by integrating the IRF over a) the spatial region over which counts were accumulated, and b) any other observables.

We plan to use XSPEC as the fitting tool, and thus will use the set of XSPEC input files. The series of count spectra, both for the source and the background (if any), is stored in PHA2 files. The DRM is broken into the ARF file, with the livetime and effective area, and the RMF file, with the energy redistribution matrix.

The event binning tool A5 will create an ARF file for each spectrum in the series with the width of the time bin (thus there will most likely be an ARF file for each spectrum). This DRM-generation tool will modify the livetime in the ARF files to account for the deadtime during each time bin, and will add the relevant effective area. Similarly, A5 will create an RMF file with the definition of the energy channels, and the DRM-generation tool will add the relevant energy redistribution matrix.

The ARF and RMF file corresponding to each spectrum in the series spanning the burst is written into the header for that spectrum.

Inputs

Parameters

- The time and position of the burst

Data:

- FT2—pointing history (extracted by U3 from D2) for the time of the burst
- FT3—the XSPEC PHA2 file with the binned data

Databases required

D3—CALDB

Outputs

- FT3—the XSPEC PHA2 file with the corresponding ARF and RMF files recorded in the headers.
- FT4--the corresponding ARF file
- FT5—the corresponding RMF files

Performance requirements

The calculations should not be computationally expensive.

Other modules required

None

Host environment

Central server

Existing counterparts

None for LAT data.

Open issues for definition or implementation

Existing counterparts

This tool will be related to U4.

Open issues for definition or implementation

The precise form of the exposure will depend on the formulation of the fitting.

5 Analysis tools

The analysis tools produce astrophysical results. In providing the input and outputs for each tool the FITS file type is indicated (e.g., FT1 for event data). The input is separated into "parameters" and "data;" the latter may be input through one of the standard FITS file types.

A1. Likelihood analysis

Date: 31 May 2002 (draft v2; added to Open Issues section; DLB multimission additions); 11/27/02 DLB

Function

This is the general-purpose analysis tool for source detection and characterization. It uses the likelihood function to fit models to observed gamma rays and the likelihood ratio test to define confidence ranges for parameters of the models. Specific applications include point-source detection, flux estimation, spectrum characterization, and source location determination. Analyses are for fixed time intervals; flux histories of sources may be obtained by likelihood analyses of data for successive time intervals with the same source model.

Some applications of the Likelihood analysis tool are necessarily interactive, such as the search for point sources in confused regions of the sky, and others can be run automatically, like an initial search for point sources or generation of source location maps for specified sources.

Joint spectral fits including data from other missions will be possible by adding to the LAT likelihood the likelihood for the data from these other missions. Intercalibration parameters will be available. The tool will accommodate the different temporal durations of the LAT and non-LAT observations.

Inputs

Parameters:

- Initial guesses for parameter values for the model. (Note that the source model will include values for all of the parameters, but within the Likelihood analysis tool these can be adjusted.)
- Constraints on parameter values (especially specification of those to be held constant). For TS maps, the ranges and increments of the variable parameters.

Data:

- FT1—event data, from U1, U2, U10, or U12
- FT7—exposure matrix calculated by the Exposure calculator U4.
- FT8—A model defined by the source model definition module U7. The source modelling module defines the parameters of the model, the region of interest on the sky and the energy range of validity.
- For multimission joint fits: count spectra, background spectra, and response functions from other missions corresponding to specified point sources in the LAT FOV.

Databases required

D3 – Instrument response functions
(Gamma-ray data and exposure are provided by U1/U8 and U4.)

Outputs

- FT8—These include the model with updated parameter values, and confidence limits, along with a specification of the data (at least file names) sufficient to be read in by the Likelihood analysis tool for subsequent analyses.

- For TS maps, a FITS file can be written. Via the image and plot display tool UI2, the map (or 1-dim cuts) can also be displayed. Coordinate limits for subsequent 'fine' maps can also be selected graphically.

The model being used for the analysis (defined by U7) can be updated, e.g., to add or remove point sources, within the Likelihood analysis tool and subsequent likelihood analyses performed.

The Likelihood analysis tool should have a provision for command logging to facilitate re-running or scripting analyses

For multimission analysis there may be additional intercalibration parameters (to accommodate inter-detector normalization errors) that are fit. The fits to the 1D count spectra from other missions can be displayed.

Performance requirements

Performance requirements will be considered in detail in the full requirements document. The most pressing requirement on performance will probably relate to the time available in the Level 2 pipeline for transient source detection.

Other modules required

Source model definition module U7

Host environment

Client computer

Existing counterparts

The EGRET likelihood analysis tool LIKE is not likely to carry over to LAT data analysis. The user interface is awkward and the interface to the instrument response functions is completely different. Also, the likelihood analysis for EGRET was designed for pointed observations, for which the instrument response functions (in particular the PSF) that apply to all gamma rays in a given region of the sky are the same. This will not be the case for LAT observations obtained in scanning mode.

Some time ago, Pat Nolan prototyped a sensible interactive interface to the LAT likelihood analysis for point sources and this should be revisited for contents and/or implementation. See http://www-glast.slac.stanford.edu/software/Workshops/January01Workshop/talks/glike.pdf

Open issues for definition or implementation

1. Should the gamma-ray data be binned before the likelihood function is evaluated? If the data are to be binned, what binning is optimal, or what is the tradeoff in sensitivity for various levels of binning? Binning might be performed on any of the parameters of the gamma rays, e.g., region of the sky of arrival, range of inclination angle (region of the sky in instrument coordinates), energy, or photon class. Unbinned analysis is the most sensitive in principle, using all of the information from each gamma-ray event, but implementation is difficult (owing to multidimensional integrations that must be performed accurately), and the loss of sensitivity if carefully chosen bins are used is likely to be small.
2. What are the implications of the Protassov et al. paper (http://arxiv.org/abs/astro-ph/0201547) about the validity of the likelihood ratio test for determining, e.g., whether or not a point source is present at a given location? At the very least, distributions of likelihood ratio values in the null hypothesis should be investigated thoroughly using simulations.
3. Should there be a facility within the likelihood analysis tool for automated analyses of the same region for multiple time ranges? This would require the exposure calculator U4 to generate multiple versions of the exposure (see Open Issues in U4), and at least a good educated guess about the proper time interval for the analyses. This kind of study will undoubtedly be done frequently. The question is whether the capability explicitly should be part of A1 and U4 or whether the scriptability of both can adequately accommodate the need.
4. How will analyses of moving sources, in particular solar system bodies, proceed? The moon is a fairly bright EGRET source. The sun is not a steady source, but is impulsively very bright. During those times, it probably could be analyzed as any point source, but more generally the question needs to be addressed of how to accommodate sources that are not fixed on the sky (but have calculable trajectories).

A2. Source identification

Date: 15 September 2002 (v1)

Function

This tool evaluates probabilities of chance positional coincidences between a specified LAT point source (defined in terms of its coordinates and the dimensions and orientation of its 95% confidence contour) and astronomical catalogs of potential counterparts. The analysis and implementation is conceived of as similar to that described by Mattox et al. (1997, ApJ, 481, 95) for quantitatively investigating the association of EGRET point sources with flat spectrum radio sources. For A2 the catalogs will not necessarily be limited to radio sources (see Open Issues, however).

Inputs

Photon data (from D1, extracted by U1, processed by U10)
Range of pulse frequencies, the first frequency derivatives, and the second frequency derivatives to be searched; can be a single ephemeris (from D4, extracted by U11) or a combination of a single ephemeris and desired ranges about the ephemeris
Choice of photon arrival times (at the satellite, at the geocenter, at the solar system barycenter, or binary-demodulated)
Energy and inclination angle-dependent cut parameters

Databases required

D5 – LAT point source catalog
D6 – includes the astronomical catalogs for counterpart searches

Outputs

Candidate counterparts among the catalogs searched, ranked in order of decreasing probability of chance occurrence.

Performance requirements

The search is not envisioned to be computationally intensive.

Other modules required

Only standard UI items and U9 catalog access tool

Host environment

Client's computer

Existing counterparts

None available as a tool

Open issues for definition or implementation

The statistical properties of the astronomical catalogs used for counterpart searches must be known very well for quantitative assessment of probabilities to be feasible. For example, the flux limit, or its distribution over the sky.

A3. Pulsar profile & periodicity tests

Date: 16 September 2002 (v3); 11/27/02 DLB (v4)

Function

This tool epoch-folds photon arrival times at trial pulse frequencies and estimates significance of pulsation at each frequency. The tool may be run to look for periodicity in any type of source. A range of pulse frequencies, the first frequency derivatives, and the second frequency derivatives will be scanned. It displays a pulse profile for a given set of those timing parameters in the ranges. When significant pulsation is found, it evaluates the timing parameters at the time of observation with their uncertainty intervals.

Inputs

Parameters:

- Choice of photon arrival times (at the satellite, at the geocenter, at the solar system barycenter, or binary-demodulated)
- Range of pulse frequencies, the first frequency derivatives, and the second frequency derivatives to be searched; can be a single ephemeris (from D4, extracted by U11) or a combination of a single ephemeris and desired ranges about the ephemeris
- Energy and inclination angle-dependent cut parameters

Data:

- FT1—event list (from D1, extracted by U1, with cuts applied by U2 to select the subset for folding to increase the pulsed fraction, processed by U10)

Databases required

None

Outputs

Plot or list of periodicity significance as a function of trial frequency and frequency derivatives

Pulse profile at a given frequency and its derivatives

Candidate ephemerides, with an evaluation of their significance and their uncertainty intervals

Performance requirements

The calculation can be computationally intensive, depending on the number of trial frequencies and frequency derivatives. For extensive searches, a period search tool (A4) may be more efficient than this tool.

Other modules required

Plotting tool, preferably with GUI front end

Host environment

Client's computer

Existing counterparts

FTOOLS XRONOS

Open issues for definition or implementation

None

A4. Pulsar period search

Date: 13 September 2002 (v4, added cuts for preferentially selecting gamma rays associated with the pulsar); 11/27/02 DLB (v5)

Function

This tool searches a photon data set for a pulsar periodicity when a pulsar ephemeris is unknown. The tool may be run to look for periodicity in any type of source. A range of pulse frequencies, the first frequency derivatives, and the second frequency derivatives will be searched.

Inputs

Parameters:

- Range of pulse frequencies, the first frequency derivatives, and the second frequency derivatives to be searched
- Choice of photon arrival times (at the satellite, at the geocenter, at the solar system barycenter, or binary-demodulated)
- Energy and inclination angle-dependent cut parameters

Data:

- FT1—event list (from D1, extracted by U1, with additional cuts applied by U2 to enhance pulsed fraction, processed by U10 when necessary)

Databases required

None

Outputs

Candidate ephemerides, with an evaluation of their significance

Performance requirements

This is likely to be a very computationally intensive tool, and extensive period searches may require CPUs more powerful than the average user's computer. The tool should estimate the time required to conduct the search before actually conducting the source.

Other modules required

Plotting tool, preferably with GUI front end

Host environment

Client's computer (may need to be run on particularly powerful CPUs)

Existing counterparts

TBD

Open issues for definition or implementation

None

A5. GRB event binning (EventBinning)

Date: 9 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tool bins GRB event data in time and energy from the LAT, GBM and/or other missions. The tool also creates a background file. The binned data may be plotted by U13 or fit by A8.

The output is in the file formats used by XSPEC. XSPEC uses 4 types of files: PHA with the number of events per time and energy bin; ARF with the livetime and effective area for each bin; and RMF with the energy redistribution matrix and the energies of the channel boundaries. The PHA2 format is a variant of the PHA format with multiple count spectra. In general this tool will bin the events into many time bins, and thus PHA2 files will be produced.

The user will be offered a variety of methods of identifying the time bins:
1. User defined bins—The user inputs the time boundaries
2. Uniform bins—The user inputs the time width for all time bins
3. S/N bins—The user inputs the desired S/N for each bin
4. Bayesian Blocks

The user must also specify the desired energy bins. For the automated time binning methods (#3, #4), the user must specify the energy band to be considered.

In some cases there will not be event data for part of the burst, but there will be binned data. This might result from the finite size of the buffers for event data. If necessary, binned data before and after the count data (e.g., for the GBM data) will be concatenated with the binned counts.

The input of energy or time bin boundaries should be possible both interactively and through files. The time bins may be derived from already binned data, e.g., the user may want to bin the LAT and GBM data using the time bins of data from another mission.

Information on estimating the background count spectrum for each time bin will be input. Since this binning tool will have multimission capabilities, there will be appropriate background-calculation code for each mission considered. Few if any LAT background photons are expected in the region around the burst during a burst's short duration, and we will most likely be able to ignore the issue of background estimation for analyzing the LAT data.

Inputs

Parameters:

- Parameters relevant to the particular method of time binning

Data:

- Energy bin boundaries
- Time grid, if necessary

- FT1--LAT event list (from D1, extracted by U1, perhaps modified by U2; a special photon cut for bright, localized sources may be used)
  Or
- Unbinned GBM count data in the standard photon FITS format
- Background information for GBM data
  Or
- Swift mask-tagged events
- Event data from other missions (e.g., AGILE, Swift) in their native FITS formats.

Databases required
None required

Outputs
- FT3—binned spectra (PHA2 format)
- FT3—corresponding background (PHA2 format), if necessary

Performance requirements
The binning is not computationally intensive nor does it require much memory.

Other modules required
None

Host environment
Client computer

Existing counterparts
Existing burst tools such as RMfit are capable of binning time-tagged events.

Status
This tool already exists as a C++ FTOOL for Swift data. The methodology for LAT and GBM data has been implemented in an IDL prototype, and will be added to the C++ tool in the near future. The "Bayesian Block" algorithm is being translated from IDL into C++.

A6. GRB rebinning

Date: 9 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tool rebins the time bins for GRB data that is already binned. The tool will only combine existing time bins into longer time bins. The tool could be applied to any binned data, but will most likely be used when only binned data exists (e.g., the GBM's "continuous" data types). The binned data may be plotted by U13 or fitted by A8.

The user will be offered a variety of methods of identifying the time bins:

1. User defined bins—The user inputs the time boundaries
2. Uniform bins—The user inputs the time width for all time bins
3. S/N bins—The user inputs the desired S/N for each bin
4. Bayesian Blocks

For the automated time binning methods (#3, #4), the user must specify the energy band to be considered.

The input of time bins should be possible both interactively and through files. The time bins may be derived from already binned data, e.g., the user may want to rebin the GBM data using the time bins of data from another mission.

Inputs

Parameters:

- Parameters relevant to the particular method of time binning

Data:

- Energy bin boundaries
- Time grid, if necessary

- FT3—LAT binned spectra (from A5)
- FT4—corresponding ARF file
  Or
- FT3—GBM binned spectra
- FT4—corresponding ARF file
- FT3—corresponding GBM background
  Or
- FT3—Swift binned spectra
- FT4—corresponding ARF file
  Or
- FT3—Binned spectra from other missions (e.g., AGILE)
- FT4—corresponding ARF file

Databases required
None

Outputs
- FT3—rebinned counts (PHA2 format)
- FT3—corresponding background file (PHA2 format)
- FT4—corresponding ARF file

Performance requirements
The binning should not be computationally intensive or require much memory.

Other modules required
None

Host environment
Client computer

Existing counterparts
Existing burst tools such as RMfit and SOAR are capable of rebinning data that were already binned in time.

Status
This tool may be combined with A5.

A7. GRB temporal analysis

Date: 9 Sept. 2002 (draft v4); 11/27/02 DLB (v5)

Function

This tool will provide the user with a variety of tools to perform standard and specialized temporal analysis methods on burst lightcurves. Included will be:
Fourier transforms
Wavelets
Cross-correlations between different energy bands
Pulse decomposition.
The tool will operate on both binned and event data.

Inputs

Data:

- FT1--LAT event list (from D1, extracted by U1, perhaps modified by U2; a special photon cut for bright, localized sources may be used)
  Or
- Unbinned GBM count data in the standard photon FITS format
- Background information for GBM data
  Or
- Swift mask-tagged events
- Event data from other missions (e.g., AGILE, Swift) in their native FITS formats.

Or

- FT3—LAT binned data from A5 (PHA2 format)

Or

- FT3—GBM binned data from either A5 or continuous data accumulated by the GBM (PHA2 format)

Or

- FT3—Data from other missions (e.g., Swift, AGILE) in XSPEC PHA2 format

Databases required

None

Outputs

- Plots and values relevant to each technique.

Performance requirements

Different techniques will vary in their resource requirements. The analyses should be able to run on the user's computer in a reasonable amount of time (e.g., less than an hour for the most computationally intensive technique operating on a large amount of data).

Other modules required

None

Host environment

Client computer

Existing counterparts

While these tools have been applied to gamma-ray burst lightcurves, they have not been incorporated into standard analysis packages such as SOAR or RMfit. These techniques are in various tools such as esearch, crosscor, and autocor.

Open issues for definition or implementation

Which techniques should be included?
Should there be separate tools for binned and unbinned data?

A8. GRB binned spectral analysis (XSPEC?)

Date: 9 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tool performs standard spectral analysis of LAT, GBM and/or other missions' binned burst spectra. The data to be fitted are assumed to be one-dimensional vectors of binned counts. Joint fits to spectra from different detectors will be feasible.

When analyzing LAT data, the LAT data must be extracted over a region ~1-2 PSFs in radius, with a correction for the burst photons (="counts" in standard burst analysis terminology) that fall outside the extraction region. U1 and U2 will extract the LAT photons from the desired region, while A5 will bin the photons in time and energy, resulting in a series of count spectra spanning the burst. U14 will create the detector response matrix (DRM), which is related to the LAT instrument response function (IRF) through an integration over the extraction region; U14 corrects for the photons that fall outside the extraction region.

The blending of burst and non-burst LAT photons within the extraction region is treated by using the event rate in this region before and after the burst (most likely this background will be 0!). Because the LAT bins may be sparsely populated, the analysis may require the Cash statistic rather than χ^2.

A8 will almost definitely be XSPEC. XSPEC currently can save the results of a single fit in an XCM file; scripting will extend this capability.

This tool will fit series of spectra. This will probably require adding a script to XSPEC, if XSPEC is A8. PHA2 files can contain a series of count spectra.

Inputs

Parameters:

- Choice of fitting spectrum
- Initial parameter values

Data:

- FT3—binned data for the LAT (PHA2 format), from A5
- FT4—corresponding ARF file, from U14
- FT5—corresponding RMF file, from U14

  Or
- FT3—binned data for the GBM (PHA2 format)
- FT4—corresponding ARF file
- FT5—corresponding RMF file

  Or

- FT3—binned data for other missions (PHA2 format)
- FT4—corresponding ARF file
- FT5—corresponding RMF file

Databases required

None

Outputs

- FT6—resulting spectral fits

Performance requirements

The fitting should be fairly fast.

Other modules required

None

Host environment

Client computer

Existing counterparts

This is the standard spectral analysis found in XSPEC, RMfit and SOAR. SHERPA can also perform these fits.

Status

XSPEC is currently unable to fit series of spectra without human intervention. This capability can be created through a script. Similarly, XSPEC does not have a standard output format for the parameters from a spectral fit; we will have to create such a format.

Open issues for definition or implementation

Will the large number of GBM counts overpower the few LAT photons in a joint fit? Is a relative weighting of the data from different detectors necessary or justified?
At what burst intensity do the assumptions underlying the standard spectral analysis break down?
If the F-test and the Maximum Likelihood Ratio Test are inappropriate for most comparisons of different model fits, what tools should be provided to assist in such comparisons? Can we provide tables appropriate for our data? Should we build in simulation tools for the user to perform the necessary Monte Carlo calculations?

A9. GRB spectral analysis tool for unbinned energy

Date: 10 Sept. 2002 (draft v5); 11/27/02 DLB (v6)

Function

This tools performs spectral analysis of LAT, GBM and/or other mission's burst data binned in time but not in apparent energy. This tool will use likelihood techniques, and thus will be related to A1. However, the analysis of burst event data will involve many fewer dimensions, permitting the use of techniques that will not be feasible with A1. In particular, analytic derivatives of the spectral model with respect to its parameters can be used in the likelihood maximization.

The LAT photons will be extracted from a region around the burst for a specified time period. Thus this tool will analyze only a single spectrum. Once the photons are extracted, the time and position associated with each photon will not be used, only the unbinned apparent energy. The general exposure tool U4 must integrate the exposure over the region from which the photons were extracted.

Inputs

Parameters:

- Initial parameter values
- Choice of fitting spectrum

Data:

- FT1—LAT event list (from D1, extracted by U1, perhaps modified by U2); a special photon cut for bright, localized sources may be used
  Or
- GBM count data
  Or
- Unbinned data from other missions (e.g., Swift, AGILE).

- FT7—Exposure data from U4 for LAT data
  Or
- Exposure data for GBM data
  Or
- Exposure data provided by the software system for other detectors

Databases required

None

Outputs

- FT6—resulting spectral fits

Performance requirements

If there are many counts then the calculation may be computationally expensive.

Other modules required

None

Host environment

Client computer

Existing counterparts

The Harvard astrophysical statistics group has written a similar tool.

Open issues for definition or implementation

Should this tool be written to do joint fits between different detectors?
Should this tool be written to analyze more than one spectrum at a time?

A10. Spectral-temporal GRB physical modeling (SHERPA?)

Date: 10 Sept. 2002 (draft v3); 11/27/02 DLB (v4)

Function

This tool fits a GRB physical model to gamma-ray burst data. The burst data may be from the LAT, GBM or other detectors; joint fits will be possible. Currently we plan to fit event data.

The tool will provide the user with a choice of physical models with adjustable parameters. A physical model with a set of parameter values will produce a burst flux that is a function of time and energy; the calculated photon flux will be folded through the instrument response functions to predict the observed events. This tool will fit the adjustable parameters by maximizing the likelihood for the observed event data. Physical models may be computationally intensive, often requiring integrals over different distributions.

At least one physical model (colliding, shocked shells) has been coded by GLAST members (Omodei et al.). Additional tools will be included. The tool will be designed so that it will be easy for a user to add his/her favorite physical models.

The choice of a physical model may be guided by spectral fits (by A8 or A9) or temporal analysis (e.g., pulse fitting in A7). Some of the model parameters may be fixed at values determined by observations at other wavelengths (e.g., a redshift). Initial values may result from spectral fits by the A8 or A9 tools.

Provision for scripting should allow re-invoking a previously performed fit.

Inputs

Parameters:

- Initial parameter values
- Choice of fitting spectrum

Data:

- FT1—LAT event list (from D1, extracted by U1, perhaps modified by U2); a special photon cut for bright, localized sources may be used
  Or
- GBM count data
  Or
- Unbinned data from other missions (e.g., Swift, AGILE).

- FT7—Exposure data from U4 for LAT data
  Or
- Exposure data for GBM data

Or

- Exposure data provided by the software system for other detectors

Databases required

None

Outputs

Fitted parameter values for the physical model
Goodness of fit

Performance requirements

Calculation of the physical model may be computationally expensive.

Other modules required

Fitting statistic maximization.

Host environment

Client computer

Existing counterparts

None.

Status

SHERPA has built-in the capability of fitting a series of spectra with a spectral model whose parameters have a temporal dependence. However, this capability has not been tested. In addition, SHERPA currently does not handle correctly the ARF and RMF files for the series of spectra. This inadequacy should be corrected in the May, '03, release of the software.

Open issues for definition or implementation

1. Should the data be binned (in time and energy) or unbinned?

6 Observation simulation

Simulated data are necessary for the development of the tools, their testing and for users who want to see what a source might look like in LAT data.

O1. Livetime/pointing simulator

Date: 15 Sept 2002 (draft v2); 11/27/02 DLB (v3)

Function

This module generates the equivalent of the output of U3, the pointing, livetime, and mode history extractor, for a simulated observation. It understands the observing strategies available and constraints on the observations, such as limitations on slewing rates, requirements for Sun-normal orientation of the solar panels, and cessation of data taking during passages through the SAA. It can be used to predict the spacecraft's pointing for tools required for supporting the mission timeline, or planning GI observations.

Inputs

Parameters:

- Orbit parameters - initial time, inclination, altitude, longitude of ascending node (standardized by an orbital geometry model)
- Observation type (survey, pointed, pointed scan)
- Observing strategy, e.g., for avoiding occultation by the earth (will have a set of standard options available)

Databases required

None LAT-specific, although the position of the sun and possibly also the moon must be calculated from the appropriate ephemeris.

Outputs

- FT2--simulated pointing/live time history. Although the simulator will be available as a tool, sets of precomputed pointing/livetime histories for various observing strategies will likely be sufficient for most users interested in generating gamma rays for simulated observations.

Performance Requirements

The spacecraft location, attitude, and the positions of the sun, earth, and moon should be reasonably accurate (~1° in celestial coordinates for the solar system bodies), and be calculated taking into account orbital precession. These simulations are not expected to be computationally intensive, and so a number of history entries corresponding to a record about every 30 seconds for a few years should be reasonable to generate.

Other Modules Required

None.

Existing Counterparts

A good orbit/attitude simulator, currently held in the astro package of the LAT instrument simulation software, is a prototype. See Open Issues.

Open Issues for Definition or Implementation

1. When the sun is near the orbital plane, rapid rolling in azimuth is implied by the sun-normal constraint for the solar panels (and the requirement to keep one side of the spacecraft cold). Various control strategies are possible for keeping the orientation of the spacecraft within the sun angle constraints without violating the constraints on slewing rate. We should determine the strategy that will be implemented for the spacecraft and use the same in O1.

O2. High-level observation simulator

Date: 24 April 2002 (draft v1); 11/27/02

Function

This module generates a list of gamma rays with their simulated as-observed parameters for a simulated observation. The output will be in the same for as the output of the Event summary data extractor (U1) and therefore suitable for analysis with any of the standard tools. This will be useful for establishing the correctness of the analysis algorithms as well as the scientific performance of the LAT and the feasibility of any proposed study.

Inputs

Parameters:

- Region of the sky of interest (potentially a subset of the coverage of the flux model)

Data:

- FT8—Flux model from U7
- FT2—Pointing/livetime/observing mode history, because the instrument response functions depend on observing mode.

Databases required

D3 (CALDB), the instrument response functions

Outputs

FT1—simulated LAT event list with observed parameters. The photon list must be trimmed according to the parameter specifications in the exposure map. The output must also include a specification of the input model, e.g., the source fluxes, positions, spectra, or at least the name of a file containing a specification of the model.

Performance Requirements

No particular requirements. Generation of simulated gamma rays at this high level is not expected to be computationally intensive.

Other Modules Required

Source model definition tool U7 and Interstellar emission model U5

Host Environment

Run on central server. (TBR)

Existing Counterparts

None.

Open Issues for Definition or Implementation

1. How should/could time dependence of source fluxes, or even periodicity of sources, be incorporated into the simulation? Time dependence of fluxes will ultimately be very

important to simulate; what is the best way to specify the parameters? (These questions are related to the implementation of U7.)

7 User interface

This section describes aspects of the user interface such as graphics and GUIs.

UI1. Event display

No requirements summary has been prepared; this is to be inherited from the LAT instrument simulation software development group

UI2. Image/Plot display

Date: 13 Sep 2002 (draft v2)

Function

This is a general-purpose plotting and image display package. It should provide platform-independent operation (among the operating systems supported), and be capable of generating PostScript output. It should accept input directly from other tools and have provision for returning graphical inputs, e.g., cursor positions. At a minimum, x-y plotting with linear and combinations of log scales should be supported, along with 2-dimensional contour maps and image displays with color bars. A selection of color tables must be provided, as well as means for specifying and saving user-defined color tables. Image scaling must also be adjustable. Overlaying of images and contours and images and plotted points should also be supported. For images, axis ranges must be specifiable and must be properly indicated in the displays.

Inputs

(for both interactive and non-interactive versions)
Data values
Plots: axis ranges, axis scaling, plotting symbol
Images: axis ranges, contour levels or color table and image scaling

Databases required

None

Outputs

A display window or a PostScript file

Performance requirements

The display generation must be quick enough for useful interactive use of the analysis tools that generate displays, but this is not likely to be a problem.

Other modules required

None

Host environment

Client computer.

Existing counterparts

There are many, here are links to a couple:
ChiPS Chandra Plotting and Image Manipulation Tool
http://cxc.harvard.edu/ciao/download/doc/chips_html_manual/index.html
PGPLOT

http://www.astro.caltech.edu/~tjp/pgplot/

Open issues for definition or implementation

None.

Image and Plotting Packages

Existing software packages shall be used for implementing the imaging and plotting services. The Imaging and Plotting Software (IPS) must meet the following requirements:

- The IPS must be freely available for use, modification, and re-distribution.
- Any external libraries used by the IPS must be well-supported, tested, and free.
- The IPS must be available on all supported platforms.
- The IPS must provide an API for C++, and optionally, for JAVA or for a high-level scripting language.
- The IPS must be extensible, allowing the creation of custom widgets.
- The IPS must be simple to install, preferably through a binary distribution for end-users.
- IPS must be able to generate publication quality images, including PostScript.
- The IPS must allow export of images into standard browser supported formats (e.g., GIF, JPEG, TIFF, etc.)
- The IPS must provide facilities to create line, scatter, and contour plots.
- Support for 3D and surface plots is desirable.
- Plots shall be modifiable, rather than forcing regeneration of plots from scratch.
- Users shall have control over plot details such as axis range, bin size, linear vs. logarithmic scaling, etc.
- The IPS must display and possibly rotate images.
- The IPS must support overlaying of images and contours and images and plotted points.
- The IPS must provide for returning interactive graphical inputs such as cursor position.
- Multiple color tables must be available, including a facility for specifying user-defined color tables.
- Color tables must be scalable, preferably via mouse or cursor inputs.
- The display rate should be less than about 2 seconds for tasks that are not data intensive.
- Interactive sessions should provide logging capabilities.

UI3. User Support

Date: 13 Sep 2002 (draft v1.2)

Purpose

This document specifies the requirements for user support.

Documentation

- User Guides and Reference Manuals: HTML and pdf versions of these documents shall be made available on the web. Both versions of each document will be kept up-to-date and consistent with each other, i.e., the HTML and pdf versions will contain the *same* information. Updated versions of the documentation will be provided with each release of the software. In each version, a log of the revision dates and summaries of the changes shall be included. All PDF versions of the documentation will be archived.
- Analysis Threads: A web page of "analysis threads" shall be provided and maintained. These are essentially very detailed use cases that show the precise sequence of commands to do a particular analysis, along with appropriate commentary and explanations. There are numerous examples on the CIAO web pages, http://asc.harvard.edu/ciao/documents threads.html.
- Script library: There shall be a web accessible library of useful analysis scripts that have been produced internally or submitted by users. Where appropriate, a script shall exist for a corresponding analysis thread. Web page access or an email address shall be provided for users to submit scripts for the on-line library. Scripts will be required to adhere to our TBD coding standards. Each script must be vetted and tested by the SSC and/or LAT software team before being accepted. All scripts will be stored in a central CVS repository to allow version tracking.

On-line Help

- man pages: For tools that are executed from the Unix/Linux shell command line, documentation in the form of man pages, similar in format to Unix-style man pages or FTOOLS' fhelp shall be provided. In addition, access to man-like information shall be provided from within the user interfaces for programs such as the Likelihood Tool. This information can be provided through methods that are part of the extension modules for each of the supported scripting languages.
- Web page help: Searchable web pages containing the information in the man pages shall also be provided. Hypertext links for all supported concepts shall be implemented. This help will also be available in a downloadable form to allow off-line access to the help facilities.

Software Installation

- Software distribution: Source code shall be available and made part of the standard software distribution package. Precompiled binaries shall be made available on the web in appropriate formats (e.g., RPMs) for all of the supported platforms.

• Installation test suite: A package of scripts and data examples that test and illustrate the use of the software shall be included with the distribution.

Troubleshooting

• FA web page: A web page providing answers to Frequently Asked Questions shall be maintained and kept up-to-date.

• Bug reporting and help-desk: A web page and help-desk email address for reporting bugs shall be provided. The format for reporting bugs along with guidelines for describing software problems, including reproducibility constraints and source-code-level characterization, shall also be given on the web page. All bug reports will be investigated within TBD amount of time. Bug tracking software will be considered to aid in documenting bugs and their ultimate resolutions. The help desk will handle users problems with the software. An initial response to all help desk questions will be provided within TBD amount of time.

UI4. Command-line interface and scripting

Date: 13 Sep 2002 (draft v1.0)

Purpose

This document specifies the requirements for the command line and scripting interfaces of the Science Tools that will have those capabilities, such as the Likelihood Tool (A1).

Commands

Each command shall perform a single, well-defined analysis task. In order to ensure "easy" compatibility with the APIs of several different scripting languages, command arguments and output shall consist only of integers, .floating point numbers (single or double precision), strings, or C/C++ pointers. Depending on the command, the termination of user input will be signified either by a newline or an end-of-line. Other data constructs such as structures, classes, or objects shall not be passed to or returned by any command. Output from the commands can be given as the return value of the function or directed as text to standard output or standard error. Error codes will adhere to a standard set of codes that are yet to be determined.

Scripting

The scripting language shall provide a means of combining common sequences of commands. However, its use should not be required to analyze the LAT data. The user should be able to perform any analysis, via a GUI for example, without having to learn the scripting language.

Command Line Interface Requirements

The capabilities of the command line interface shall include:
1. Command recall and completion (e.g., GNU readline)
2. Session and error logging
3. System call access
4. Online help (via a man-like facility or launchable GUI)
5. Interrupt handling
6. Session recovery
7. Configurability, via environment variables and a user-supplied startup script or as specified in a predefined default configuration. The environment variables and script will allow the user to specify such things as preferred coordinate systems, directory paths, log file names, etc..

Scripting Language Requirements

The scripting language must have the following qualities and/or capabilities:
1. Flow control (e.g., loops, conditional constructs)
2. Math operations (trigonometric and special functions, etc.)
3. Allows dynamic loading of modules
4. Extendable/Embeddable

5. Provides extensions for GUI programming
6. Supports object oriented programming
7. Programmable exception handling

In the coming months, the above requirements will be used to evaluate the scripting language options. The candidate languages that have so far been considered are Perl, Python, Tcl, and Ruby. All of these languages satisfy requirements 1–5 and 9, and the most current versions of these languages support object-oriented programming.

UI5. GUI interface and Web access

Date: 11 Sep 2002 (draft v0.1)

Function

This module will develop the requirements for GUI development that will exist for the benefit of the science tools. A plan for achieving portability and consistency across platforms should be determined. Draft requirements for GUI development libraries and tools. Evaluate and recommend tools to use to develop the GUI interface. Develop a schedule for GUI prototypes and testing.

GUI Requirements

A graphical user interface shall be provided for all analysis tools and utilities. A standard scheme for specifying "functional look-and-feel" features shall be implemented and applied uniformly across all software applications and platforms. Such features shall include, but not be limited to:

Location and organization of menus, menu items, buttons, and taskbars.

Availability and implementation of tasks such as "Print...", "Save", "Save as...", "Help...", and "Exit".

Location, size, and shape of windows used for plot and image display, command confirmation (e.g., "Really run /bin/rm –rf /*?"), and error messages.

Methods of data entry.

The visual appearance (e.g., colors, appearance of buttons, fonts, etc.) of all the GUIs shall be implemented consistently on each platform for all the applications. However, uniform visual appearance across platforms and/or "native look-and-feel" shall not be required.

A single GUI shall be implemented from which all of the command-line tools may be run. All functionality available at the command line interface shall also be accessible via the GUI. This includes opening plotting windows, specifying datasets and search paths, executing scripts, etc. This GUI will include:

A menu of all command-line (i.e., non-interactive) tools to provide an access point for running each tool. The menus shall be organized hierarchically, grouping related tools by function. Because some command-line tools are naturally run in specific sequences, options shall be exist for writing intermediate data to files and/or for piping those data directly to the next tool.

Menu directives for launching the interactive analysis tools (e.g., the likelihood analysis tool A1), in either command-line or GUI mode.

Configuration Parameters for Command-line and Interactive Tools

Users will be provided with a mechanism to control input parameters and configuration settings. The GUI will include a menu option to allow users to control default parameters and configuration settings. Parameter values will have persistence throughout a single session, i.e. when returning to a data entry point, the parameter values from the most recent visit will have been retained.

The default configuration settings will be stored in the same file(s) used by the tools at the command-line. The tools shall be run using the parameters appearing in the configuration menu, but the files storing the default parameters must not be changed until the user issues an explicit command to save the updated parameters (via a "Save" button on the configuration window or a "Save parameters" menu item). Upon ending a GUI session, the user shall be queried if she wishes to save the current configuration if it has not already been saved.

Widget Set/Tool Kit Requirements

A standard library (such as FOX, Gtk++, or Tk) will be used to develop the GUI. The library must be freely available on our supported platforms. The library must provide APIs for C/C++ and our supported scripting languages. Stability and technical support for the library will be among the key considerations when the selection is ultimately made.

Web Interface Requirements

Web interfaces to the science tools may be provided (TBR). The extent of the use of the Web as an analysis interface is yet to be determined. The most likely candidates for interfacing over the Web are the event summary database (D1) and the pointing and livetime history database (D2), which will generally accessed from remote servers.

When web access is provided, there will be support for Netscape and Internet Explorer. The layout of web pages will be standardized, consistent and intuitive. Text character size will be determined by the user's browser settings. Standard web options such as, tabbing to the next field, will be provided. When authentication and authorization are required, such as perhaps to implement the 'qualified investigator' data access restrictions for U1 and U3, a single sign-on will grant access for all services. A sensible timeout mechanism will also be implemented, requiring re-authorization after an extended period of inactivity.